Cloning and Co-Expression of hG-CSF and hSCF in *E.coli*

Sagarika Dash and Srimeenakshi S

Dept. of Biotechnology and Medical Engineering
National Institute of Technology Rourkela

Cloning and Co-Expression of hG-CSF and hSCF in *E.coli*

Dissertation submitted in partial fulfilment

of the requirements of the degree of

Master of Technology

in

Biotechnology

by

Sagarika Dash

(Roll Number: 215BM2007)

&

Srimeenakshi S

(Roll Number: 215BM2011)

based on research carried out

under the supervision of

Prof. Mukesh Kumar Gupta

May, 2017

Department of Biotechnology and Medical Engineering

National Institute of Technology Rourkela

Dept. of Biotechnology and Medical Engineering
National Institute of Technology Rourkela

Dr. Mukesh Kumar Gupta

Associate Professor

May 30, 2017

Supervisor's Certificate

This is to certify that the work presented in the dissertation entitled **"*Cloning and Co-Expression of hG-CSF and hSCF in E.coli*"** submitted by **"*Sagarika Dash*", *Roll Number: 215BM2007 and* "*Srimeenakshi S*", *Roll Number: 215BM2011 ,*** is a record of original work carried out by her under my guidance and supervision in the partial fulfilment of the requirements for the degree of ***Master of Technology*** in ***Biotechnology*** at the ***Department of Biotechnology and Medical Engineering***, ***National Institute of Technology Rourkela***. Neither this dissertation nor any part of it has been submitted earlier for any degree or diploma to any institute or university in India or abroad.

Mukesh Kumar Gupta

Declaration of Originality

We, *Sagarika Dash*, Roll Number *215BM2007* and *Srimeenakshi S*, Roll Number *215BM2011* hereby declare that this dissertation entitled *Cloning and Co-Expression of hG-CSF and hSCF in E.coli* presents our original work carried out as a master students of NIT Rourkela and, to the best of my knowledge, contains no material previously published or written by another person, nor any material presented by us for the award of any degree or diploma of NIT Rourkela or any other institution. Any contribution made to this research by others, with whom we have worked at NIT Rourkela or elsewhere, is explicitly acknowledged in the dissertation. Works of other authors cited in this dissertation have been duly acknowledged under the sections "Reference" or "Bibliography". We have also submitted our original research records to the scrutiny committee for evaluation of our dissertation.

We are fully aware that in case of any non-compliance detected in future, the Senate of NIT Rourkela may withdraw the degree awarded to me on the basis of the present dissertation.

May 30, 2017
NIT Rourkela

Sagarika Dash
Srimeenakshi S

Acknowledgement

Thinking and Feeling appreciation signifies that we believe in someone's existence rather than the divine who rules everything. This thesis entitled "*Cloning and Co-Expression of hG-CSF and hSCF in E.coli*" is the successful outcome of guidance and support bestowed on us throughout our work. We would like to express our most sincere gratitude to every single being those devoted their time for successful completion of our thesis.

We would like to show my foremost gratitude towards my Guide **Prof. Mukesh Kumar Gupta** who helped me with my thesis in every possible stage and encouraged me to do better. We would also like to thank Ph.D. scholar **Bijayalakmi Sahoo,** and Senior Research Fellow **Dr. Rakesh Bhaskar** and **all our lab mates** for their painstaking effort for the completion the thesis.

Finally, we would not miss to thank heartily to our parents and close friends for their unconditional love and encouragement shown towards me and completion of the thesis. Last but not the least, we would like to thank the Department of Biotechnology and Medical Engineering for providing such a stimulating atmosphere and nice work environment.

May 30, 2017
NIT, Rourkela

Sagarika Dash
Roll Number: 215BM2007

Srimeenakshi S
Roll Number: 215BM2011

Abstract

Granulocyte Colony Stimulating Factor (G-CSF) is a hematopoietic growth factor that controls the proliferation, differentiation and the function of neutrophils. G-CSF is clinically used in human for the treatment of neutropenia in diseases such as AIDS, aplastic anaemia, myelodysplastic syndrome, and congenital or chemotherapy-induced neutropenia. However, the bioactivity and stability of commercially available recombinant hG-CSF such as Filgrastim and Lenograstim are lower than endogenous G-CSF. The objective of the present study is to express recombinant hG-CSF in E.coli along with SCF as a fusion partner to improve the bioactivity. In the present study, in silico analyses were performed for finding the possible mutant variants of G-CSF which may increase the stability of the recombinant hG-CSF for its incorporation into hG-CSF by site-directed mutagenesis and also analyse the binding affinity of the G-CSF-SCF fusion protein with G-CSF receptor (G-CSF-R) and SCF receptor (SCF-R). A total of 5 mutant variants of hG-CSF was generated and docked with GCSF-R using Hex dock 8.0 software. In order to analyse the binding affinity of the G-CSF-SCF fusion protein, docking analyses of the fusion protein with GCSF-R and SCF-R were performed using patch dock server. Human G-CSF gene (547 bp) was isolated from human Umbilical Cord Blood and U-87 cell line. The hG-CSF gene was cloned into TOPO-TA vector for sequencing followed by cloning into the pET14b vector for expression using Nde I and Bam HI restriction ends. Human GCSF gene was then end modified to fuse with the fusion partner. Besides, Human SCF gene (567 bp) was purchased from GenScript™ in the pUC57 cloning vector. It was restriction digested from the pUC57 vector using Bam HI and Xho I restriction enzymes. The restriction digested hG-CSF, SCF inserts were inserted into pET14b vector containing Nde I and Xho I restriction ends. The ligated expression vectors were then transformed into chemically competent DH5α cells followed by plasmid extraction and transformation into expression host E.coli BL21 (DE3).The expression of human G-CSF and G-CSF-SCF protein in E.coli was confirmed using SDS-PAGE analysis. Further expression profile of the proteins was optimised to increase the protein expression. From the in silico analysis, it was found that the mutant variant 5 may have improved biostability than the wild type G-CSF variant. The study was also found that the G-CSF-SCF fusion protein has high binding affinity to G-CSF-R and SCF-R from

the global energy values. Furthermore, increased level of G-CSF and G-CSF-SCF fusion protein observed under optimised IPTG concentration of 1mM, post-induction duration of 8h and at 3% of ethanol concentration.

Key words: *G-CSF, SCF, Fusion protein, In silico analysis, BL21DE3, cloning, transformation, expression, optimization.*

Contents

Supervisor's Certificate i
Declaration of Originality ii
Acknowledgement iii
Abstract iv
List of Figures viii
List of Tables x
Chapter 1 1
Introduction 1
Chapter 2 6
Literature Review 6
2.1 Human granulocyte colony stimulating factor 6
2.1.1 The structure of G-CSF 6
2.1.2 The genetics of G-CSF 8
2.1.3 The functions of G-CSF 8
2.1.4 Clinical Importance of G-CSF 8
2.2 Hematopoietic growth factors 9
2.3 G-CSF fusion protein 10
2.4 Stem Cell Factor (SCF) 10
2.4.1 Biology of Stem Cell Factor 11
2.5 Synergistic effect of G-CSF and SCF 11
Chapter 3 15
Objectives 15
3.1 Objectives 15
Chapter 4 16
Material and Methods 16
4.1 Preparation of Blood Cell Pellet for Extracting RNA 16
4.2 Total RNA Extraction Protocol: RNA-XPress^{TM} Reagent 16
4.3 First strand cDNA synthesis 17
4.4 Polymerase Chain Reaction for amplification of human G-CSF 18
4.5 Agarose gel electrophoresis 19
4.6 Restriction Digestion 19

4.7 Gel elution of DNA 20
4.9 Colony PCR for screening transformants 21
4.10 Recombinant Selection and Plasmid Isolation 21
4.11 Quantification of Isolated DNA 21
4.12 Ligation of DNA fragments 21
4.13 Preparation of E. coli competent cells 22
4.14 Transformation of E. coli cells 22
4.15 SDS-Polyacrylamide gel electrophoresis 22
4.16 Coomassie brilliant blue staining 23
4.17 Bioprocess and Protein Expression 23
4.18 *In silico* Analysis of hG-CSF with mutations and fusion partner for improved bioactivity 23
Chapter 5 25
Results and Discussion 25
5.1 Isolation of human G-CSF from U-87 human glioblastoma cells 25
5.2 Isolation of human G-CSF from UCB 26
5.3 TOPO TA cloning of human G-CSF 27
5.4 Cloning of human G-CSF in pUC57 cloning vector 28
5.4 Isolation and restriction digestion of pET14b expression vector 29
5.4 Ligation of human G-CSF with pET14b expression vector 30
5.5 Cloning of pET14b-hGCSF construct in DH5α 31
5.6 Cloning of pET14b-hGCSF construct in BL21DE3 32
5.7 *In silico* analysis of hG-CSF with mutations and fusion partner for improved bioactivity 33
5.8 Cloning of SCF in pUC57 cloning vector 37
5.9 Ligation of pET14b-hGCSF-SCF 38
5.10 Cloning of pET14b-hGCSF-SCF in BL21DE3 39
5.11 Expression of human G-CSF 40
5.12 Expression of human G-CSF-SCF 41
Chapter 6 43
Conclusion 43
References 45

List of Figures

Figure 1: Crystal structure of recombinant human G-CSF..... 2
Figure 2: The effect of multiple hematopoietic growth factors in the production of neutrophil and megakaryocytes [35] 3
Figure 3: Control Elements in pET Vector System 4
Figure 4: BL21DE3 strain 5
Figure 5: Native structure of hG-CSF protein [27] 7
Figure 6: G-CSF and G-CSF Receptor interaction. A: Homo-dimerization B: G-CSF-R activation..... 7
Figure 7: Alternative splicing of G-CSF mRNA [38] 8
Figure 8: Hematopoietic Growth Factors [50] 9
Figure 9: Integration of SCF and G-CSF signaling pathways [21]. 12
Figure 10: Phase contrast microscopic images of U-87 human glioblastoma cells Magnification: 10X. Scale bar: 100 µm 25
Figure 23: Restriction digestion of pUC57-GCSF was performed with NdeI and and BamHI endonucleases. 29
Figure 24: Vector Map of pET14b 29
Figure 27: Insert and vector in agarose gel. Lane 1: Insert(hGCSF),2: Vector(pET14b) . 30
Figure 28: Transformation plates with ligated vector and insert..... 31
Figure 29: PCR verification of ligated colonies. Lane 1,3: PCR with hG-CSF specific primers.2:1kb Ladder..... 31
Figure 30: Transformation plates of pET14b-hGCSF construct in BL21DE3 A: Transformation plate B: Control Plate..... 32
Figure 31: PCR verification of transformed colonies. Lane 1,2,3,4,6,7,8: PCR with GCSF specific primers.5:1kb Ladder 32
Figure 32: Proteins modeled using SWISS-MODEL and visualized using PyMOL. A: Native protein B:M1 mutant C:M2 mutant D:M3 mutant E:M4 mutant F:M5 mutant 33
Figure 33: Receptor Protein interaction using Hex 8.0.0.A: Native protein B:M1 mutant C:M2mutant D:M3 mutant E:M4 mutantF:M5 mutant..... 34
Figure 34: Sequence and schematic model which has shown the construct of GCSF and SCF bound together by the SGGGGSGGGGSGGGGS linker. 35
Figure 35: Graphical representation of secondary elements in chimeric GCSF-SCF protein..... 35
Figure 36: I-Tasser server was used to predict the 3D model of GCSF-SCF fusion protein 35
Figure 37: Docking of GCSF-SCF with GCSFR using patchdock 36
Figure 38: Docking of GCSF-SCF with GCSFR using patchdock. To examine the protein ligand interactions, the models for ligand binding potency has predicted. 36
Figure 39: Transformation plate of pUC57-SCF to DH5α Competent Cells A: Transformation plate B: Negative control plate 37
Figure 42: Restriction Digestion of pET14b with Nde I & Xho I. Lane 1:1kb Ladder 2: Restriction digested GCSF product 3: Restriction digested SCF product..... 38
Figure 46: Transformation of plasmid to BL21DE3 competent cells A: Transformation plate B: Negative control plate 39
Figure 47: Growth curve and specific growth rate curve of induced and uninduced culture

..... 40
Figure 48: Growth curve at different ethanol concentration. 40
Figure 49: SDS-PAGE result of induced and uninduced culture of GCSF..... 41
Figure 50: SDS-PAGE result of induced and uninduced culture of GCSF-SCF 41

List of Tables

Table 1:cDNA reaction composition 17
Table 2: cDNA Synthesis Conditions 17
Table 3: PCR reaction composition 18
Table 4: PCR programme 18
Table 5: Primer Designed for GCSF expression 19
Table 6: RE Digestion Reaction Composition 19
Table 7:TOPO Cloning Reaction 20
Table 8:Colony PCR 21
Table 9: ETOTAL values for the various interactions 34

Chapter 1

Introduction

Therapeutic proteins are produced from their native source in a very small amount and so for its use in the pharmaceutical market necessitates a large scale production. The principles of genetic engineering can be therefore use for the production of recombinant proteins to meet the market demand. Large scale production of recombinant proteins for their application in the field of biotechnology and medicine includes cloning a gene of interest in a desired expression vector under the control of an inducible promoter and subsequent expression of it. The first expression platform of recombinant therapeutic proteins was the Escherichia coli system. E. coli is a well-established host that offers easy genetic manipulation, short culturing time, and low cost. Furthermore, Engineered E. coli strains has ability to produce a wide variety of different types of proteins includes protein containing disulfide bonds. With a large knowledge on the molecular genetics, simple cultivation requirements such as growth on inexpensive carbon sources, rapid biomass accumulation and a short generation time *E. coli* remains the first choice as an expression host [12].

Colony Stimulating Factors (CSF) are the group of cytokines involving in the hematopoiesis process formation in which a small number of self-renewing stem cells proliferate and differentiate into mature blood cells. Different CSFs act on bone marrow cells to generate different lineages of hematopoietic cells. Granulocyte colony stimulating factor (G-CSF) is a hematopoietic growth factor that stimulates the proliferation, differentiation and the function of neutrophils. Cancer patients receiving chemotherapy are vulnerable to potentially life-threatening infections due to suppression of white blood cells production. G-CSF thus has been permitted for the treatment of neutropenia in cancer patients, a condition which appears after chemotherapy. Furthermore, G-CSF also has neuroprotective properties which have been used as a protective agent in a variety of neurodegenerative diseases.

Human granulocyte colony-stimulating factor (G-CSF) is an 18.8 kDa protein. Alternative splicing of G-CSF mRNA encodes two transcript variants of 207 and 204 amino acid residues respectively. The initial 30 amino acid residues at N-terminal region, responsible for secreting the mature protein outside the cell. The insertion or deletion of 3 amino acid residues around the 35th amino acid residue is responsible for the difference between two transcript variants of G-CSF.

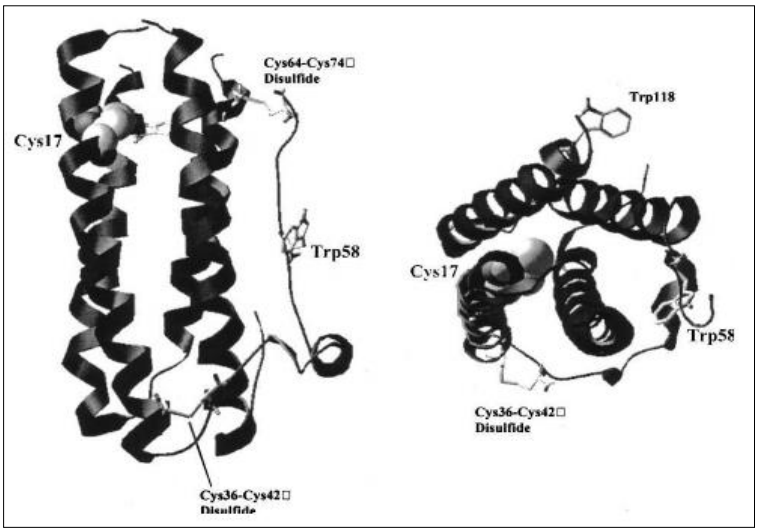


Figure 1: Crystal structure of recombinant human G-CSF.

Transcript variant 2 with 174 amino acid residues is found to be 20 times more active than the transcript variant 1 with 177 amino acids. The nucleotide sequence of G-CSF contains two intramolecular disulphide bonds between Cys36-Cys42 and Cys64-Cys74 and a cysteine at residue 17. Furthermore, O-glycosylation site at Thr133 in G-CSF is not essential for biological activity. Accordingly, the biological activity of nonglycosylated recombinant hG-CSF is as same as glycosylated hG-CSF. So the G-CSF can be effectively produce using prokaryotic protein expression system.

The hematopoietic system is a hierarchical structure in which non-replicative mature blood cells are synthesized by multipotent stem cells. Several hematopoietic growth factors play a major role in hematopoiesis. Some growth factors (e.g., GMCSF) have a broad array of effect on very early hematopoietic progenitors, leading to multilineage increases in hematopoietic cell differentiation, while others (e.g., G-CSF) appear to act

mainly on more terminally differentiated cell types, producing moderately profound changes in specifically committed populations, such as neutrophils.

Hematopoietic growth factors have redundant properties. Redundancy is defined as the production of the same effect by different factors. In addition, the factors are pleiotropic, implying that different effects can be mediated by one cytokine. Some cytokines such as GMCSF, SCF, IL-3, IL-6, and IL-1 1 have been proved to stimulate granulopoiesis with efficiencies less than that of G-CSF. But it has also been shown that if those cytokines act synergistically with G-CSF, maximum granulopoiesis can be attained. One way to increase the effect of G-CSF in granulopoiesis is the combined administration of G-CSF with other hematopoietic growth factors. Several finding indicates that G-CSF exert it activities by acting synergistically and sequentially with other cytokines mainly GMCSF, SCF, IL-3 which raises the potential application of G-CSF fusion proteins. Stem Cell Factor is one of the cytokines that plays a major role in hemetopoiesis, melanogenesis and spermatogenesis. If the synergistic effect occurs between G-CSF and SCF, fusing both of them will increase the bioactivity of the growth factor.

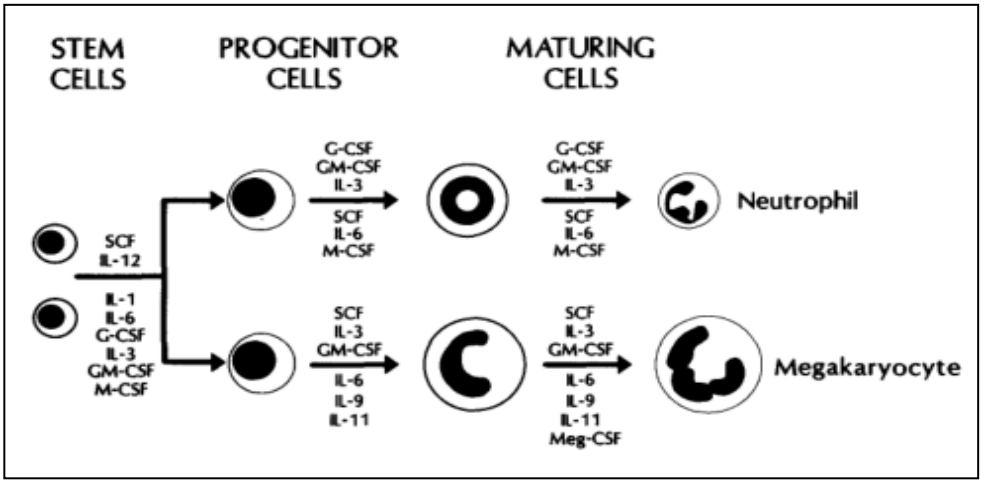


Figure 2: The effect of multiple hematopoietic growth factors in the production of neutrophil and megakaryocytes [35]

E.coli T7 expression system is the most reliable and successful system for the production of heterologous proteins like Insulin and human growth hormone. In T7 expression system, a gene of interest is cloned downstream to the T7 promoter of the expression vector which is then transformed into T7 expression host. T7 expression hosts carry the

phage T7 polymerase gene which upon the addition of inducer results in the transcription of the inserted gene of interest. Also high density culture achieved with simple process scale up makes the culture process economical and so this platform has been used for successful production of recombinant proteins like insulin and human growth hormone.

The pET vector System is the most effective system for the expression and cloning of recombinant proteins in *E. coli*. pET vector system has been the potent choice for the expression of proteins that have been difficult to produce using *E. coli* promoter-based systems such as *tac*, $p_{L,}$ *lac*, have been successfully cloned and expressed stably in the pET System. Target gene expression in pET plasmids is induced by T7 RNA polymerase of the host cell. Apart from the bacteriophage-based promoter, pET vectors also carry several of the fusion tags at the 5' end. Besides, pET vectors also contain protease cleavage sites such as thrombin, enterokinase that provides selective removal of tags following purification.

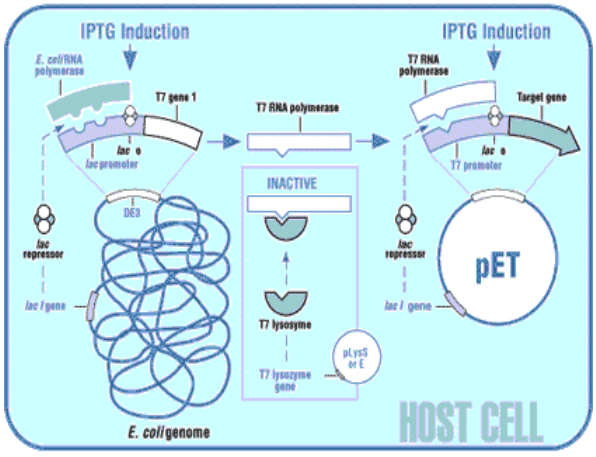


Figure 3: Control Elements in pET Vector System

E. coli BL21 (DE3) is an ideal strain for the T7 promoter based expression systems like pET and , pRSET . BL21DE3 contains the lambda DE3 lysogen. DE3 expresses T7 RNA polymerase from the lacUV5 promoter to allow simple induction of recombinant proteins with IPTG. Expression of non-toxic recombinant proteins in E. coli is higher in BL21 (DE3) cells than in BL21 (DE3) pLysE or BL21 (DE3) pLysS.

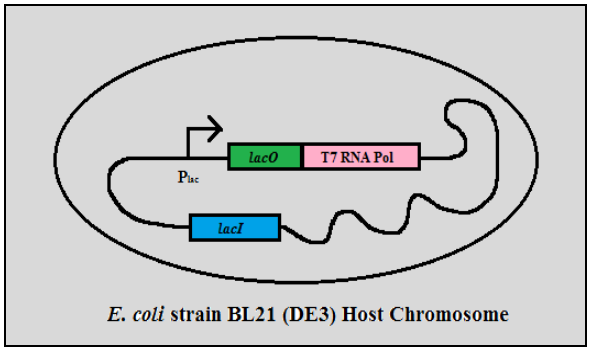


Figure 4: BL21DE3 strain

But the basal level expression of heterologous genes are considerably higher in BL21 (DE3) pLysS or BL21 (DE3) pLysE. However, a shortcoming associated with the use of *E. coli* system is the inability to synthesize recombinant proteins with post translational modification cells (e.g., human, Chinese hamster Ovary) has been successfully used as expression systems.

In silico analysis of the interaction of wild and mutated proteins with the G-CSF receptor and SCF receptor can give us an inference of how the wild and mutants are interacting with the receptor by calculating the binding energy values. Negative energy values predict better interaction of the ligand and receptor. Protein modelling software can be used for modelling the mutant proteins which can be then docked with the receptor using a docking software. Mutations created can change the stability of the protein. So to evaluate the change in stability upon mutation various protein stability determining software can be used. For achieving maximum production of a product the process condition for growth of the bacteria must be optimized.

Chapter 2

Literature Review

Impressive advancement in recombinant DNA technology in the past decades has produced several therapeutic proteins for clinical application. Scientists are intended to better the technology to speed up the clinical trial process as there are thousands of therapeutic proteins are lined up for the approval. [13]. One of the major hurdles in the application of therapeutic proteins in clinical applications is the reduced half-life and activity of the therapeutic protein compared to its wild type counterparts. Expression system employing gene fusion constructs is the possible solution to increase the bioactivity and stability of the therapeutic protein. The advent of synthetic biology also made it possible to express chimeric fusion proteins to increase the overall bio activity of the therapeutic protein. One of the greatest outcomes of expressing fusion protein remains in the field of biopharming. Biopharming is the production of physiologically important therapeutic proteins through genetically engineered plants and animals [29] in high volume.

2.1 Human granulocyte colony stimulating factor

Human granulocyte colony-stimulating factor (hG-CSF) is a hematopoietic growth factor involves in controlling differentiation, proliferation, and efficient activation of blood cells [48]. It is an important growth factor for neutrophil proliferation both in vivo and in vitro. Large quantities of recombinant human G-CSF are produced using genetic engineering and it is successfully used to neutropenia condition in cancer patients which appear after chemotherapy.

2.1.1 The structure of G-CSF

The native hG-CSF is an 18.9 kilo Dalton glycoprotein consisting of 174 amino acid residues. It has O-linked carbohydrate group linked to Thr 133 [38, 39, 40]. The X-ray crystallography studies have revealed that the three-dimensional structure of G-CSF contains four antiparallel alpha helical structures with the left-handed twist [5]. The four helices are represented as helices A-D. The loops linking the helices are showed as AB, BC, and CD (Figure 5). Human G-CSF contains two disulphide bonds at positions Cys36-

Cys42 and Cys64-Cys74.Human G-CSF contains two disulphide bond at positions Cys36-Cys42 and Cys64-Cys74.

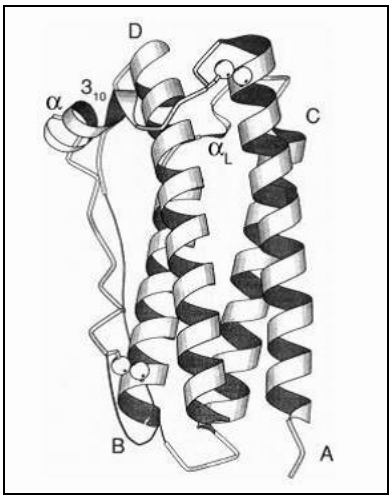


Figure 5: Native structure of hG-CSF protein [27]

G-CSF shares four alpha helix bundle with two long crossover connection structure with GMCSF, GH, IFN- α, IL-2, and IL-4. Although hG-CSF is an important protein not only for basic science but also for therapeutic use, the detailed activation mechanism of hG-CSF-R has not been clarified till yet [46]. Moreover, study of the crystal structure has revealed the atoms involved in the molecular positioning and recognition for the 2:2 complex; there is ligand-receptor interaction at two locations (site II and III). Thus, homodimerization of receptors by this 2:2 complex is elucidated to be an important feature for hG-CSF-R activation.

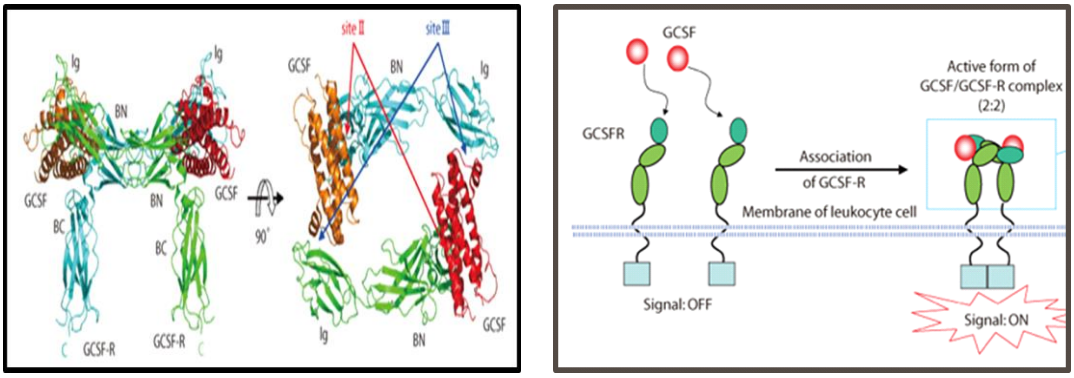


Figure 6: G-CSF and G-CSF Receptor interaction. A: Homo-dimerization B: G-CSF-R activation

2.1.2 The genetics of G-CSF

Gene for human G-CSF is present in the long arm of chromosome 17. G-CSF gene contains 5 exons and 4 interrupting introns. Primary G-CSF transcripts give two different transcript variants upon splicing. As shown in the figure 7, two sequences for alternative splicing are organized in tandtem at the 5' end of intron 2. Transcript variant 1 is formed by splicing at 380/381 position whereas transcript variant 2 produced by splicing at 371/372 position. The second donor sequence [38]. More than 80% of hG-CSF mRNA produced belongs to transcript variant 2 type.

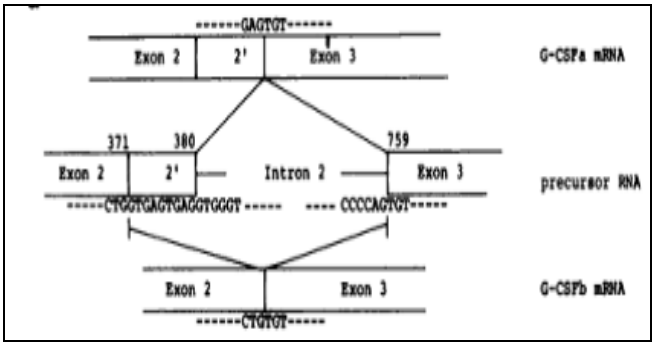


Figure 7: Alternative splicing of G-CSF mRNA [38]

2.1.3 The functions of G-CSF

Due to increased occurrences of bone disorders, the need for bone grafts is increasing corresponding to the demand. But, the availability of the autografts and allografts depends on the various factors which cannot be controlled. After several years of research still, the autograft serves as the gold standard as it does not face Graft rejection from the host's body, the biggest disadvantages are double surgery and highly expensive. Allografts are highly histocompatible they can also be used after decellularization of bone matrix, bone chips etc.[10]. The various other strategies adopted to repair the bone disorders or fractures include bone cement and involvement of bone morphogenetic proteins. The bone graft designed in vitro should be osteoinductive as well as angiogenic in nature to support the regeneration of the bone on the site of implantation. This leads to the necessity to develop the Bone grafts in vitro for implantation into the patients.

2.1.4 Clinical Importance of G-CSF

The role of G-CSF to infection associated with bacterial pathogens like Listeria

monocytogenes, Pseudomonas aeruginosa and Candida albicans has been elucidated by various researchers. L. monocytogenes causes infection by propagating within the macrophages and the important role of G-CSF in suppressing bacterial load by promoting phagocytosis has been well studied in vivo using G-CSF deficient mice models. G-CSF plays a major role in bacterial clearance by controlling macrophage responses and circulating neutrophil levels.

2.2 Hematopoietic growth factors

The significance importance of the bone marrow microenvironment relies mainly on the production of hematopoietic growth factors that are important for the proliferation, differentiation and eventual function of monocyte, granulocyte, basophils and eosinophils. Erythropoietin, produced by the interstitial cells in kidney, is responsible for the production of RBC.

FIGURE 2: HEMATOPOIETIC GROWTH FACTORS

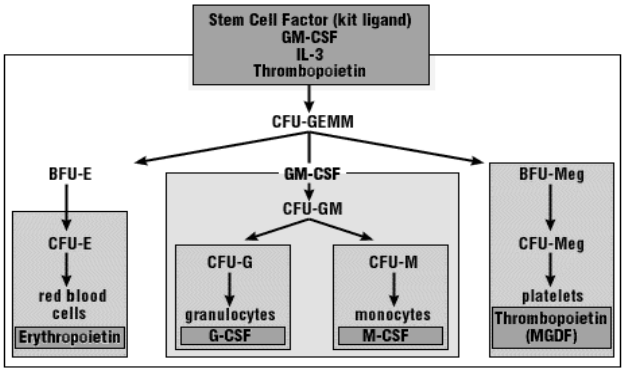


Figure 8: Hematopoietic Growth Factors [50]

Furthermore, early acting growth factors such as GMCSF, SCF, flt3 support initial growth of hematopoietic cell line. Furthermore, GMCSF also stimulates the phagocytic and cytotoxic activity of monocyte and granulocyte. G-CSF, on the other hand, stimulates the functional characteristic of granulocytes. MCSF increases the functional activity of monocytes. Sequential and synergistic interaction of different growth factors is important for the differentiation of hematopoietic cell line. It has been reported that the in vivo combination of EPO, G-CSF, GMCSF, SCF and IL-11. The in vivo effects of a combined administration and all possible combinations of erythropoietin (EPO), G-CSF, interleukin 1 (IL-11), and stem cell factor (SCF) have been studied and reported by modifying the quantitative and qualitative dose response characteristics of the cytokines [50].

2.3 G-CSF fusion protein

Several methods have been used to enhance the half-life of drugs. Increasing the half-life of protein and peptide drugs by protein fusion technology is one of the most widely used methods in molecular biology practices. Based on the development of molecular biology and genetic engineering, some natural proteins with long half-life have been used as fusion partners to enhance the circulating half-life of drugs, such as IgG-Fc, transferring (tf), and human serum albumin (HSA). HSA-G-CSF fusion protein has been studied and reported that it has longer half-life than the native G-CSF protein [53]. The G-CSF-tf fusion protein has also been reported to increase the absolute neutrophil count upon oral administration than the 3 native G-CSF [7]. The bacterial stress responsive proteins such as bacterioferritin (Bfr), chemotaxis protein cheZ (CheZ), HTH-type transcripttional regulator yjdC (YjdC), HTH-type transcripttional regulator yjdC (YjdC), HTH-type transcripttional regulator yjdC (YjdC), peptidyl-prolyl cis–trans isomerase B (PpiB), and glutathione synthetase (GshB) have been used as a fusion expression partner with G-CSF and reported that these stress responsive proteins act as a folding enhancer. Similarly, high yield and high solubility of G-CSF protein have been reported by combining Maltose Binding Protein (MBP) and Protein Disulphide Isomerases (PDI) with G-CSF [21]. G-CSF/SCF fusion protein with higher biological activity than native G-CSF has been also patented (WO2015047062 A1, 2013) The present study aims to produce fusion protein of G-CSF with hematopoietic growth factor such as IL-3, GMCSF and SCF and to increase the biological activity of the protein.

2.4 Stem Cell Factor (SCF)

Stem Cell Factor is a co stimulatory Hematopoietic Growth Factor plays a major role in proliferation and differentiation of the early hematopoietic stem cells. SCF, when combine with cytokines such as G-CSF, GMCSF, IL-3, TPO, EPO, resulting in a synergistic improvement of the development and survival of hematopoietic cells. On marrow cells in semi-solid culture, SCF shows pronounced synergy with factors such as IL-6, IL-1, IL-3, IL-7, granulocyte CSF, granulocyte-macrophage CSF, and erythropoietin [2, 26,33, 54, 55]. The colonies generated are larger and more numerous than those with the latter factors individually. Apart from Hematopoiesis, SCF also have an important role in melanogenesis and spermatogenesis. The therapeutic applications of

SCF consist of hematopoietic stem cell (HSC) mobilization, gene therapy, immunotherapy and ex vivo stem/progenitor cell expansion. It appears that SCF enhances the proliferation and differentiation of progenitors in these cultures but that lineage commitment is controlled by the later-acting factor [4].

2.4.1 Biology of Stem Cell Factor

In human, gene code for SCF protein is present at the long arm of chromosome no 12 whereas in mouse, it is located at the SI locus of the chromosome [2, 6, 26, 54, 55]. Alternative splicing of SCF results in the soluble and transmembrane forms of SCF that either includes or excludes a proteolytic cleavage site [11]. The proteolytic cleavage occurs after Ala 165. The transmembrane form of SCF (SCF 220) lacks exon 6 in which amino acids 149-177 are substituted by a Gly residue. In its native state, SCF exists as a monomer. SCF undergoes spontaneous dissociation–reassociation of monomers to form the non-covalent SCF homodimer. The specific activity of SCF dimers are 10–20 times higher than the specific activity of the monomer. On purified human progenitor stem cell populations in vitro, it has a unique maintenance-enhancing capability and when given in combination with other cytokines, such as granulocyte CSF, granulocyte-macrophage CSF, IL-6, IL-3, and erythropoietin, there is marked proliferation which is not seen with the other factors given alone or in combinations (WO2015047062 A1).

2.5 Synergistic effect of G-CSF and SCF

Hematopoietic cell maturation and function depends on the action of several cytokines and growth factors. These growth factors modulate the differentiation of the progenitor cells and controls the final stage cell functions. Synergistic effect of SCF with G-CSF plays a significance role in hematopoiesis. SCF and G-CSF are important growth factors for the growth of human primitive hematopoietic cells in vitro [10]. In addition, the increased mobilization of blood progenitor cells (PBPC) was found to be higher with the combination of SCF and G-CSF in vivo than that was with G-CSF alone [3, 10, 18]. Furthermore, several studies have reported that the SCF and G-CSF combination increases the CD34 cells counts in patients at risk of poor PBPC mobilization [22, 37, 44].

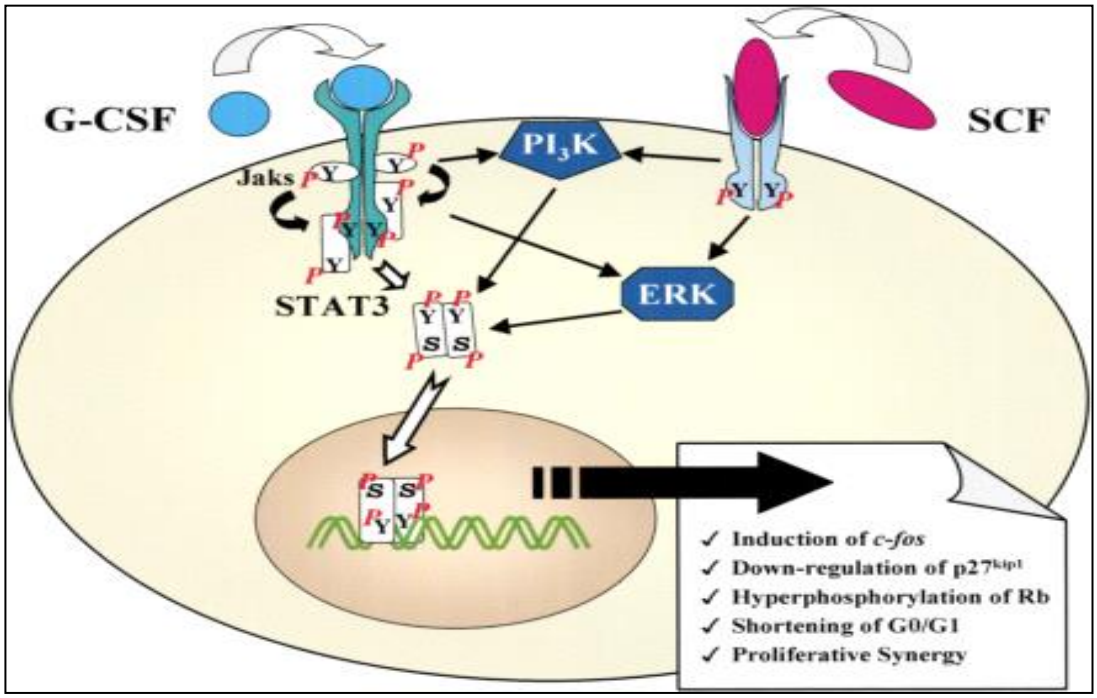


Figure 9: Integration of SCF and G-CSF signaling pathways [21].

Signaling pathways of G-CSF and SCF upon binding with its respective receptors, unite on STAT-3signal transduction, which become phosphorylated on serine and tyrosine residues, and carry out the consequent biological effects such as MAPK, Jak–STAT and PI3K signaling pathways [21].

2.6 Recombinant forms of G-CSF

Similar to the other colony stimulating factors, G-CSF is also present in small quantity in our system similarly. The quantity of CSFs has been reported as low in macrophages, fibroblasts, monocytes, endothelial cells, and bone marrow cells [34]. In order to use G-CSF as a drug, the gene of G-CSF was cloned and expressed as a recombinant protein [42], [43]. Currently, recombinant hG-CSF (rhG-CSF) is a recognized commercially available recombinant protein, has been used mainly for neutropenias. There are several forms of recombinant G-CSF is currently available such as filgrastim, neupogen, nartograstim and pegfilgrastim. Opti-nartogratimhas been successfully produced and reported as a codon optimized form of (for bacterial expression system) nartograstim which shows similar biological activity as native G-CSF [32]. In this present study we are trying to produce the G-CSF fusion protein in BL21DE3 host. Filgrastim, a recombinant

form of G-CSF is commercially being used as a drug to treat neutropenia condition after chemotherapy, and in patients undergoing bone marrow transplantation. Filgrastim a non-glycosylated recombinant human G-CSF has a methionine group at its N-terminal end [43]. Lenograstim a glycosylated recombinant human G-CSF) is produced using Chinese Hamster Ovary cells with 4 % glycosylation. Pegfilgrastim is the non-glycosylated human G-CSF with PEG moiety to mask the proteolytic cleavage sites and reduces the renal excretion of the protein thereby elevating the serum level of G-CSF [43]. Nartograstim, a mutant variation has also been expressed in E. coli with mutations at N-terminus and a substitution of cysteine at its 17^{th} position [25].

2.6 Engineering human G-CSF for improved bioactivity

Various recombinant preparations of human G-CSF has been engineered with enhanced biological activity and physiochemical properties. Site directed mutagenesis, a protein engineering tool has been used for modification of specific sites in a nucleotide sequence for production of mutants. The simplest and most widely used approach in this regard is the Quick Change Site-Directed Mutagenesis System developed by Stratagene. Using this approach, desired mutations can be introduced using a thermal cycler and with a pair of complementary primers containing the desired mutation. G-CSF has one free cysteine at its 17^{th} aminoacid position and has two disulfide bridges (Cys36-Cys42 and Cys64-Cys74). It has been reported in various literatures that free cysteine molecules are involved in intermolecular thiol/disulfide interchange that leads to formation of inactive oligomers. The Cysl7 was substituted with aminoacids like Gly, Ser, Tyr, Arg, Ala, He, and Pro, or was deleted using site-directed mutagenesis in order to study the effect on thermo stability and bioactivity of the protein. The results obtained inferred that the replacement of Cysl7 with Ala improved the heat-stability of mutated protein by five times than the wild type [28]. KW-2228(Neu-up) is a modified human granulocyte colony-stimulating factor in market with enhanced bioactivity and stability and is known to possess more granulopoiesis activity than the wild-type human G-CSF. In the wild type G-CSF Cys^{17} is situated in the core surrounded by hydrophobic residues. When Cys^{17} was substituted with Ser, hydrogen bond was formed in between hydroxyl group of Ser and carbonyl group of Leu^{14} that eventually lead to the stability of the protein. Similar results were obtained for all other N terminal mutations [24]. Amino acid residues in the G-CSF receptor binding interface has been mutated to construct rHSA/G-CSF mutants. Mutant

G-CSF fusion protein is a fusion protein that stimulates the proliferation of neutrophilic granulocytes, and has a basic structure of G-CSF conjugated with carrier protein or carrier protein with G-CSF, in which the G-CSF has one of the K34H, L35I, K40H, and L41I substitutions or its combinations [51]. The mutants designed were expressed in Pichia pastoris and for the obtained mutants, the ligand-receptor affinity was determined by using surface plasma resonance (SPR) technology. Through the SPR technology an association rate constant and a disassociation rate constant was obtained, and a ligand-receptor equilibrium disassociation constant was determined. The larger the equilibrium disassociation constant the lower is the ligand-receptor affinity. When rHSA/G-CSF was compared with rHSA/G-CSF mutants it was found that the substitution in the sites of K34H, L35I, K40H, and L41I gave higher equilibrium disassociation constant. So compared with rHSA/G-CSF, the mutants designed had a longer half-life, lower binding affinity and an improved biological activity [51]. Variant forms of G-CSF have been designed with substitution of Leu in place of Met elucidating increased stability against oxidation [31].

G-CSF has been successfully engineered for enhancing the thermodynamic stability without interfering with the receptor binding sites. An unusual abundance of glycine residues is found in the helices of G-CSF that results in low intrinsic alpha-helical propensity [9]. Glycine to alanine substitutions is expected to increase protein stability. After experimental replacement of glycine residues with amino acid alanine at specific positions 26, 28, 149, and 150 resulted in high increase in the stability of protein. Also after evaluating first generation mutant, second and third generation mutants were also designed and studied for increased stability over wild type G-CSF. Second and third generation Gly/Ala substitutions which elucidated that with an increase in intrinsic alpha-helix propensities increased the overall protein stability [9].

Chapter 3

Objectives

3.1 Objectives

The major objective of this study was to express recombinant hG-CSF in *E.coli* expression vector system along with SCF as fusion partner to improve the bioactivity. The specific Objectives are,

1. To isolate and clone human G-CSF and SCF.
2. To construct expression vector for expression of G-CSF and SCF
3. To co-express hG-CSF and hSCF as a fusion partner
4. To optimize bioprocess parameter for expression of hG-CSF.

Chapter 4

Material and Methods

All materials were purchased from HiMedia Lab (Mumbai, India) unless otherwise specifically indicated.

Total RNA was extracted from discarded Umbilical Cord matrix and/or blood by informed consent from IGH, Rourkela and U-87 human glioblastoma cells and hG-CSF (CSF3) was PCR amplified. The hG-CSF cDNA was obtained by extracting the total RNA from Umbilical Cord Blood followed by 2 step RT-PCR. The primers were designed accordingly for both verification and cloning.

4.1 Preparation of Blood Cell Pellet for Extracting RNA

The blood was collected in appropriate Haemobag with heparin and antibiotics and was stored at mentioned temperature (4°C). The blood was transferred to 15 ml centrifuge tube with proper sterilization procedure. For 1ml of blood,3 ml of 1X of RBC lysis buffer was added. The contents were mixed properly by inverting the tubes and was allowed to stand at RT for 10min. The cells were centrifuged at 600xg for 10 minutes at RT to obtain cell pellet. The supernatant after centrifugation was decanted and the pellet obtained was gently resuspended in 1 ml of 1X RBC lysis buffer and was allowed to stand for 5 minutes at RT. The cells were pelleted for 2 minutes by centrifuging at RT at 3000 rpm. The supernatant was decanted and the pellet was carefully resuspended in 1 ml of sterile 1X DPBS. The obtained cells were pelleted at 3000rpm for 2 minutes at RT and the collected pellet was used for RNA extraction.

4.2 Total RNA Extraction Protocol: RNA-XPress™ Reagent

The cell pellet (50-100mg) was transferred to a clean nuclease free epi Tube. About 1ml of RNA Xpress Reagent was added per 50-100 mg of tissue/pellet. Homogenized & Incubated for 5 mins @ RT. Centrifugation was done at 12000g at 4°C for 10 minutes. The obtained supernatant was collected in a sterile centrifuge tube (this step was for avoiding fat and lipid bilayer mixing with the aqueous phase).200µL of chloroform (Himedia MB109) was added to it and shaken vigorously for 15 sec. The tubes were then incubated at room temperature for 10 minutes. Centrifugation was carried out at 12000g

for 15 minutes at a temperature of 4°C. The obtained supernatant was collected in a sterile centrifuge tube. About 500µL of isopropyl alcohol (Himedia MB063) was added. The tubes were inverted few times to mix briefly and were incubated for 10 min at RT. Centrifugation was done at 12000g at at 4°C for 10 minutes and supernatant was completely removed by inverting the tube upside down. 1mL of 75% ethanol (Himedia MB106) (MB Grade, in DEPC water) was added. Centrifugation was carried out at 7500 g at 4°C for 5 minutes and the pellet was allowed to air dry. The pellet was finally resuspended in nuclease free water (~50µL) and was allowed to dissolve by heating for 10 min at 55-60°C. The dissolved RNA was processed immediately.

4.3 First strand cDNA synthesis

Complementary DNA was synthesized using the *High-Capacity cDNA Reverse Transcription Kit* (Applied BiosystemsTM – 4368814) in a Veriti® 96 well thermal cyclers (Applied Biosystems) using the manufacturer's cycle conditions with modifications and random primers(supplied) as mentioned in the product protocol. The synthesized cDNA was then verified for genomic DNA contamination using the gene specific exon-exon boundary primers followed by agarose gel electrophoresis. The cDNA was then stored at -20°C. The cDNA reaction composition is given in table 1 and the PCR conditions are given in table 2.

Table 1:cDNA reaction composition

Initial Conc	Sample	Volume for 1 reaction(µl)
10X	RT buffer	2.0
25X	DNTPs	0.8
10X	Random primer	2.0
	RNA	10.0
	Reverse Transcriptase	1.0
	Nuclease Free water	4.2
	Total volume	20

Table 2: cDNA Synthesis Conditions

Step 1	Step 2	Step 3	Step 4
25℃	37℃	84℃	4℃
10 min	120 min	5min	∞

4.4 Polymerase Chain Reaction for amplification of human G-CSF

PCR was carried out to amplify the hG-CSF gene from DNA extracted from U-87 cell line and human umbilical cord blood using specific primer pairs and AmpliTaq Gold 360® Master Mix (Applied Biosystems – 4398881) in a Veriti® 96 well thermal cyclers (Applied Biosystems). The primer was synthesized using Primer 3 software, Genome Compiler and was validated using Net Primer. The specificity of the primer pairs was tested using BLASTn. The primer synthesizing was outsourced from Bioserve, Hyderabad. The PCR cycle conditions and parameters used for an amplifying specific set of genes mentioned below.

Table 3: PCR reaction composition

Initial Conc.	**Sample**	**Final Conc.**	**Volume for 1 reaction(µl)**
2X	Master mix	1X	12.5
10uM	Forward primer	0.5uM	1.25
10uM	Reverse primer	0.5uM	1.25
	DNA	60 ng/ µl	3
	GC enhancer		2.5
	DEPC water		4.5
	Total reaction volume		25

Table 4: PCR programme

Holding	Denaturation	Annealing	Extension	Final extension	4℃
95℃	95℃	62℃	72℃	72℃	∞
7 min	30 sec	30 sec	90 sec	7 min	
1 cycle		30 cycle		1 cycle	

The primer pairs used for the amplification and verification of the genes are given below.

Table 5: Primer Designed for GCSF expression		
Primer	**Forward(5'-3')**	**Reverse(5'-3')**
Human G-CSF specific primer	CAGTGCAGCATATGACC CCCCTGGGCCCT	GGGGATCCTCAGG GCTGGGCAAG
SCF specific primer	TAATGAGGATCCGGTGG AGGAGGTTCTGG	TATTATCTCGAGTCA GGCTGCAACAGG
Internal Primer	GGAGAAGCTGGTGAGTG AGT	TTCCCAGTTCTTCCATC TGC

4.5 Agarose gel electrophoresis

0.8 % agarose was prepared in 1X TAE buffer and ethidium bromide was added to it. The solution was poured to a casting tray with comb of appropriate size to set the gel. The casting tray was transferred to a gel tank filled with the electrophoresis buffer after the gel sets. DNA samples were mixed carefully with 6X DNA loading buffer and loaded onto the wells and the apparatus was runned. The DNA bands were visualized under 260 nm UV trans-illuminator.

4.6 Restriction Digestion

The restriction enzymes were purchased from New England Bio labs. Restriction digestion was performed for a time interval of 15 minutes at 37°C for the vector and the insert (hG-CSF). The RE digestion was carried out in a 25µL reaction with 5 units of restriction enzyme, 0.5µg of DNA and 1X CutSmart® Buffer each (Table 6). For all restriction digestion of pUC57-hGCSF and SCF 1 hour of incubation was followed to obtain all the restriction digested products.

Table 6: RE Digestion Reaction Composition		
Component	**Concentration**	**Volume used (µL)**
Nde I	5 Units	0.25
BamHI	5 Units	0.25
DNA	0.5µg	10
CutSmart Buffer	1X	2.5
MB Grade Water	remaining	12.5
Total reaction volume		25

After restriction digestion the samples were run in a 0.8% Agarose gel (Low EEO) (Himedia MB002). The bands corresponding to the insert and vector size was excised using a sterile scalpel blade no.11 (Himedia LA769). The RE digested gene fragments were extracted from the gel using) and the OD values were taken to ascertain the purity and concentration of the gel purified nucleic acids.

For all other restriction digestion carried out the above mentioned conditions were maintained.

4.7 Gel elution of DNA

PureLink Quick Gel Extraction Kit (Invitrogen) was used to elute the specific DNA fragemts from agarose gels. A small area of gel containing specific DNA fragment was excised and weighed. L3 Buffer was added to the gel containing DNA band in ratio of 3:1 and heated in water bath.1 volume of isopropanol was added and the mixture was loaded on to an extraction column and centrifuged at maximum speed for 1 minute. 500 µl of wash buffer was then added and centifuged at maximum speed for 1 minute. Washing step was repeated and 50 µl of the given elution buffer was then added and kept for incubation for around 10 minutes. It was then centifuged at maximum speed for 1 minute and the sample was stored at -20°C till further use.

4.8 TOPO TA Cloning and Transformation

The following table describes the TOPO® Cloning reaction.

Table 7:TOPO Cloning Reaction

Reagents required	Volume(in µl)
PCR product(fresh)	4
Salt Solution	1
TOPO® vector	1
Total Volume	6

2 µL of prepared TOPO reaction was added to chemically competent E. coli cells and mixed gently by tapping. The cell mixture was kept on ice for 15 minutes and then the cells were given heat-shock for 90 seconds at 42°C.The tubes were then transferred to ice and kept for 5 minutes. S.O.C. medium was added and the cell mixture was kept in shaker at 200 rpm at 37°C for 40 minutes.100 µL was spreaded from on a LB-Amp agar plate and incubated overnight at 37°C.

4.9 Colony PCR for screening transformants

A positive colony was carefully picked from an LB agar plate using tip of a pipette and transferred to a 0.5 ml centrifuge tube containing 50 µl of DEPC treated water. The solution was then vortexed. Then the tube was kept in a heat block at 99°C for 5 min. The tube was centrifuged at 12,000 × g for 1 minute. 5 µl of the supernatant was transferred to a 0.5 ml centrifuge tube and proceeded for PCR.

Table 8:Colony PCR

Reagents	Volume(in µl)
2xMastermix	12.5
Forward Primer(0.5µM)	1.25
Reverse Primer(0.5µM)	1.25
DNA (30 ng/µl)	5
DEPC water	5
Total Volume	25

4.10 Recombinant Selection and Plasmid Isolation

Colonies were inoculated in fresh LB broth and kept for overnight in shaker incubator. O.D. value of overnight grown culture was calculated and plasmid was isolated by both manual and miniprep method using *Charge Switch Pro Plasmid Miniprep Kit*. Isolated plasmid was runned in agarose gel and visualized under UV.

4.11 Quantification of Isolated DNA

The concentration of DNA was quantified spectrophotometrically in UV range. O.D. value of isolated DNA was measured at 230 nm,260 nm,280 nm and 320nm. The ratio of OD260 /OD280 was determined to check the purity of the sample. The concentration of DNA was calculated by using the following formula.

DNA(ng//µl) =OD260 X50 X Dilution Factor

4.12 Ligation of DNA fragments

The DNA fragments digested with restriction enzyme so as to produce cohesive end were mixed with the digested vector. Vector (4.5Kb) was mixed with the insert (0.5Kb) in a molar ratio of 3:1.50 ng of vector and 16 ng of insert was used along with 0.5 U of T4 DNA ligase(Invitogen) for carrying out ligation. The final volume of reaction was kept 20

µl and the reaction mixture was incubated at 16 °C for 16 hours. The reaction was then heat inactivated at 65 °C for 20 minutes and digested with XhoI (killer cut) prior to transformation to chemically competent cells.

For a three-way ligation all the components were maintained in the molar ratio of 1:1:1 and all other reaction conditions were kept same.

4.13 Preparation of E. coli competent cells

100 µl of glycerol stock of E. coli cells was inoculated in 5 ml of LB broth and kept in shker for overnight incubation.500 µl of the overnight grown culture was used as inoculum for a 50 ml LB broth. When the cells reached to an OD of 0.4, they were transferred to pre chilled centrifuge tube under sterile conditions. Cells were harvested by centrifuging at 3000g for 15 minutes at 4°C.Supernatant was decanted and the pellet was dissolved in 100 mL of ice cold $MgCl_2$.Cells were harvested at 2000g for 15 minutes at 4°C. Supernatant was discarded and the pellet was dissolved in 200 mL of ice cold $CaCl_2$.The cell suspension was incubated on ice for 20 minutes. Cells were then harvested by centrifuging at 2000g for 15 minutes at 4°C. Supernatant was discarded and the pellet was dissolved in 50 mL of ice cold 85 mM $CaCl_2$, 15% glycerol. Cells were harvested at 1000g for 15 minutes at 4°C.Transfer the suspension to the 50 mL conical tube. Supernatant was discarded and the pellet was dissolved in 5 mL of ice cold 85 mM $CaCl_2$, 15% glycerol. 50 µl aliquots were made in prechilled 1.5 mL centrifuge fuge tubes and snap freezed with liquid nitrogen. Frozen cells were stored -80°C freezer till further use.

4.14 Transformation of E. coli cells

100 µl of chemically competent cell was thawed on ice followed by addition of plasmid DNA. The cell mixture was kept on ice for 15 minutes. The cells were then given heat shock by keeping them in a preset water bath, at 42 ºC for 90 seconds. The cells were immediately transferred onto ice and incubated on ice for 5 minutes.SOC broth medium was added to the cell mixture kept on incubator shaker set at 37°C for 40 minutes.100 µl of cell mixture was plated on a LB ampicillin agar plate and plates were incubated at 37°C for overnight.

4.15 SDS-Polyacrylamide gel electrophoresis

Glass plates, comb and spacers were cleaned and assembled. Separating(resolving) gel from the gap between the glass plates until filled to ¾th followed by with

immediate addition of isopropanol. The gel was left up to 30 minutes for polymerization. After solidification, stacking gel was poured and immediately a comb was inserted into the stacking gel solution, taking care to avoid trapping air bubbles. Now stacking gel was allowed to get polymerize. Now the glass plate was fixed in the electrophoresis apparatus and the gel tank was filled with 1X SDS Running Gel Buffer. Samples to be separated by SDS-PAGE were prepared as follows: 5 µl of 4 X SDS loading buffer was added to 15 µl of protein sample and the mixture was heated for 10 min at 70°C.Samples were loaded into different wells. The gel was runned until the dye front reached the bottom of the gel.

4.16 Coomassie brilliant blue staining

SDS-PAGE gels containing more than 200 ng protein concentration were visualised by Coomassie Brilliant Blue R 250 (CBB R 250) staining solution that comprises of 0.1% (w/v) CBB dissolved in 25% (v/v) of methanol and 10% (v/v) of acetic acid in water. De-staining solution was prepared in the composition of water, methanol and acetic acid in the ratio of 50:40:10.

4.17 Bioprocess and Protein Expression

IPTG (1M) stock was prepared. BL21DE3-pET14b-hGCSF was cultured and induced with 1mM IPTG at 0.5 O.D. Hourly culture was harvested and O.D.(600nm) was recorded for both induced and uninduced culture. Specific growth rate for induced and uninduced culture was calculated. Graphs were plotted for growth curve and specific growth curve. After 0,4,6,8 hr 5ml of culture was withdrawn and centrifuged to obtain the cell pellet. Lysis buffer was prepared and 500µl of it was added to each cell pellet and the samples were then sonicated for 5minutes and kept at -20 ºC further use for running SDS PAGE. Similarly, IPTG concentration was varied from 0.2 to 1mM and protein expression was studied in presence of 1%,2% and 3% of ethanol. BL21DE3-pET14b-hGCSF-SCF was cultured and induced with 1mM IPTG at 0.5 O.D. Hourly culture was harvested and O.D.(600nm) was recorded for both induced and uninduced culture. Protein expression was studied for different IPTG concentration and different period of induction.

4.18 *In silico* Analysis of hG-CSF with mutations and fusion partner for improved bioactivity

Genome compiler was used at all the simulation steps of the work. SWISS-MODEL was

used to model the native and mutant proteins based on the protein sequences. PyMOL was used to visualize the models designed. Hex 8.0.0 was used for docking of proteins with the receptor. ENCoM server was used to find the ddG values due to mutations and finally RAMPAGE was used for generating Ramachandran plots for the proteins.

In silico analysis of GCSF fusion protein was carried out to measure the interaction of GCSF fusion protein with its respective receptor and to determine the hindrance of the fusion protein to GCSF. Protein sequence of GCSF, SCF, IL3, HLA, GCSF receptor and SCF receptor was retrieved from the NCBI (National Centre for Biotechnology Information) site. Sequence of fusion protein was made by combining the sequence of GCSF with its respective fusion partner via linker sequence. The protein sequence was then submitted in I-TASSER server to model the 3D structure of the GCSF fusion protein. Interaction and the binding affinity of the GCSF fusion proteins was analyzed using patch dock server.

Chapter 5

Results and Discussion

5.1 Isolation of human G-CSF from U-87 human glioblastoma cells

U-87 human glioblastoma cells were cultured and maintained. Figure 10 shows the characteristic morphology of U-87 cells as seen under phase contrast microscope.

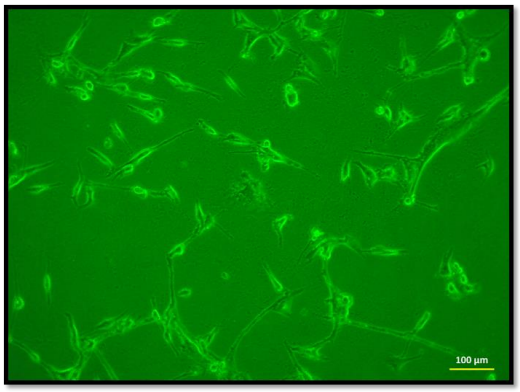


Figure 10: Phase contrast microscopic images of U-87 human glioblastoma cells Magnification: 10X. Scale bar: 100 µm

RNA was isolated from U-87 cells and converted to cDNA as shown in figure 11 and 12. Bands were obtained for mRNA and smear was observed in cDNA gel. This cDNA was amplified with hG-CSF specific FP and RP primer and hG-CSF specific EEB primers as shown in figure 13.

Band of 547 bp was obtained which inferred that hG-CSF was amplified using hG-CSF specific primers as inferred from the *in silico* analysis using genome compiler. Also a band of around 270 bp was obtained using hG-CSF specific internal EEB primers. hG-CSF obtained in the gel was cut and gel purified and runned on agarose gel as shown in figure 14. Band size of around 547bp was obtained which confirms the presence of human G-CSF.

Figure 11: RNA of U-87 Cell. Lane 1: Total RNA

Figure 12: cDNA of U-87 cell. Lane 1: cDNA

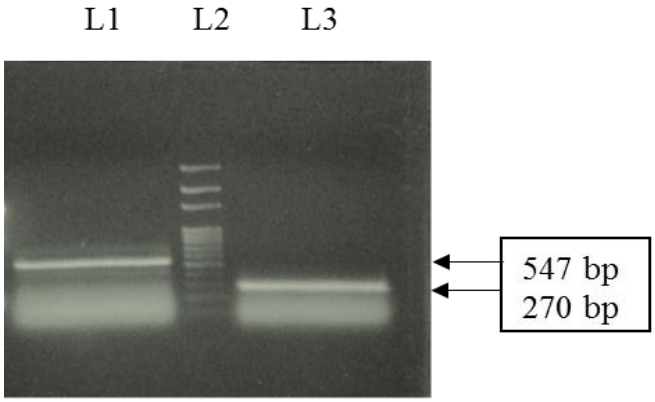


Figure 13: PCR amplification of cDNA. Lane 1: PCR product with GCSF specific primers.2:100 bp Ladder 3: PCR product with EEB primers

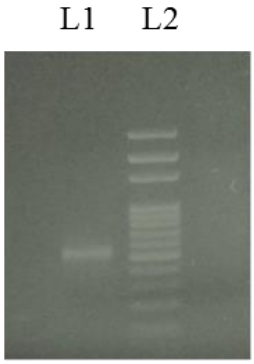


Figure 14: Gel purified GCSF. Lane 1: Purified GCSF 2:100 bp ladder

5.2 Isolation of human G-CSF from UCB

Total RNA of Umbilical cord blood was isolated from the blood pellet as shown in figure15. RNA was converted to cDNA and amplified using hG-CSF specific primers as shown in figure 16 and 17 respectively.

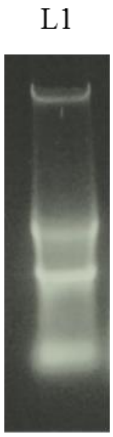


Figure 15: RNA of UCB. Lane 1: Total RNA

Figure 16: cDNA of UCB. Lane 1: cDNA

Figure 17: PCR amplification of cDNA. Lane 1,2,3: PCR product with GCSF specific primers.

Appropriate bands were obtained signifying the successful isolation of human GCSF.

5.3 TOPO TA cloning of human G-CSF

After successfully isolating human GCSF it was restriction digested and cloned in a TOPO vector. TOPO TA Cloning of GCSF was performed. Figure 18 illustrates the transformation result of the vector to the E. coli DH5α cells. The cells were then inoculated in LB broth and cultured. From grown culture plasmid was isolated as shown in figure 19.

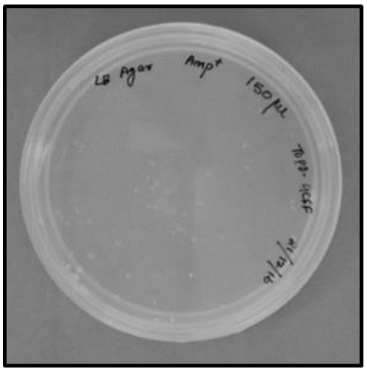


Figure 18: Transformation of pTOPO-hGCSF

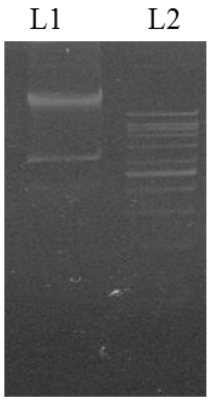


Figure 19: Plasmid Isolation. Lane 1: pTOPO-hGCSF.2:1kb ladder

Colonies were observed in the transformation plate. Band of around 4500 bp size was obtained which confirmed the presence of pTOPO-GCSF plasmid in the cells.

5.4 Cloning of human G-CSF in pUC57 cloning vector

The pUC57-GCSF plasmid was transformed to HiPuRA DH5α Competent cells. Colonies were observed in the transformation plate as shown in the figure 20. No colonies were observed in the negative transformation plate.

After successful transformation, colonies were inoculated and grown overnight to proceed for isolation of plasmid (using Charge Switch Pro Mini Prep Kit). The isolated plasmids were runned in agarose gel as shown in figure 21. The bands of appropriate size were obtained. For further verification of successful transformation, PCR Amplification of the isolated plasmid with hG-CSF specific primers was carried out as shown in figure 22.

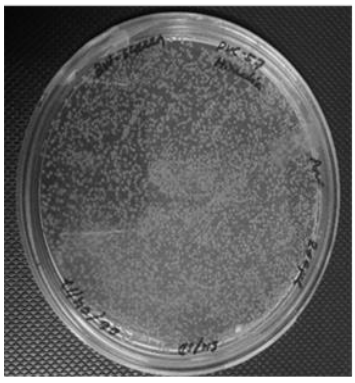


Figure 20: Transformation plate of pUC57-hG-CSF to DH5α Competent Cells

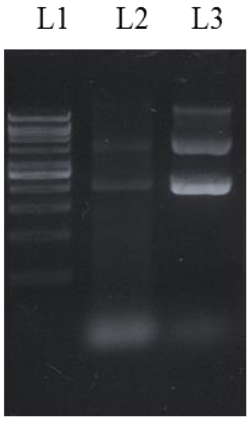


Figure 21: Isolated Plasmid of. pUC57-hG-CSF Lane 1:1kb Ladder,2,3: Isolated plasmid

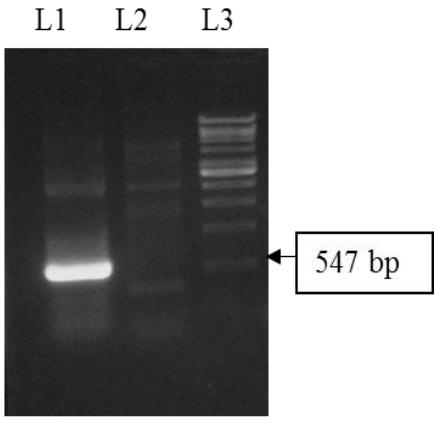


Figure 22: PCR amplification of human G-CSF with specific and nonspecific primers.

Band of size of around 547 bp was observed in the lane loaded with PCR using hG-CSF specific primers. A faint band of around 2000bp was also observed suggesting that of the remaining plasmid DNA. However, no band was observed in the PCR reaction with non-specific bands. Restriction digestion of pUC57-GCSF was performed with NdeI and and BamHI endonucleases. The agarose gel image of it is shown in the figure 23

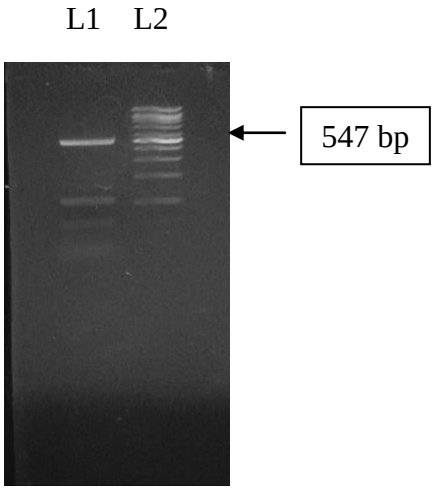


Figure 23: Restriction digestion of pUC57-GCSF was performed with NdeI and and BamHI endonucleases. Lane 1:1kb Ladder,2: Restriction digested sample

Restriction digestion of pUC57-GCSF resulted in four specific bands. The band with size of around 547bp was confirmed to be hG-CSF and thus was gel extracted for ligation with the expression vector.

5.4 Isolation and restriction digestion of pET14b expression vector

Expression vector pET14b was chosen for the expression of GCSF based on all the suitable features required for expression of the protein. The vector map is shown in figure 24.

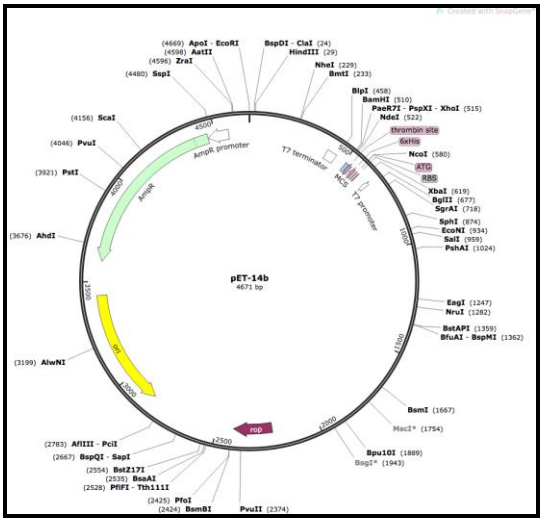


Figure 24: Vector Map of pET14b

Transformation of pET14b was carried out and the colonies obtained were grown overnight for isolation of plasmid. Specific band of size 4.6kb was obtained inferring that the isolated plasmids are pET14b as shown in figure 25. Restriction Digestion of pET14b with Nde I and Bam HI was performed and runned in agarose gel as shown in figure 26. The restriction digested sample gave a band of size was purified and proceeded for ligation with hG-CSF.

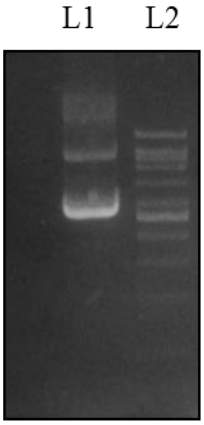


Figure 25: Plasmid Isolation of pET14b. Lane 1: pET14b.2:1 kb Ladder

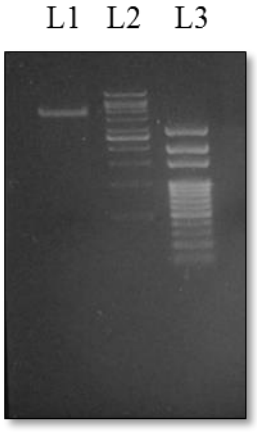


Figure 26: Restriction Digestion of pET14b.Lane1: Digested sample.2:1 kb Ladder,3:100 bp Ladder

5.4 Ligation of human G-CSF with pET14b expression vector

Restriction digested pET14b and G-CSF was purified using *Pure Link Quick Gel Extraction kit*(Invitrogen) as shown in the figure 27 and proceeded for ligation.

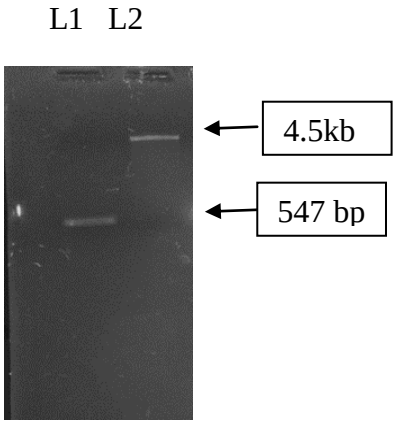


Figure 27: Insert and vector in agarose gel. Lane 1: Insert(hGCSF),2: Vector(pET14b)

Ligation of cohesive ends of human G-CSF and pET14b was performed using T4DNA ligase(Invitrogen) at 16ºC for 16 hours. Heat inactivation and killer cut (using XhoI) was done to increase the specificity of ligation and transformation.

5.5 Cloning of pET14b-hGCSF construct in DH5α

Cloning of ligated pET14b-Hgcsf was performed in competent cells. The transformed plates are shown in the figure 28.

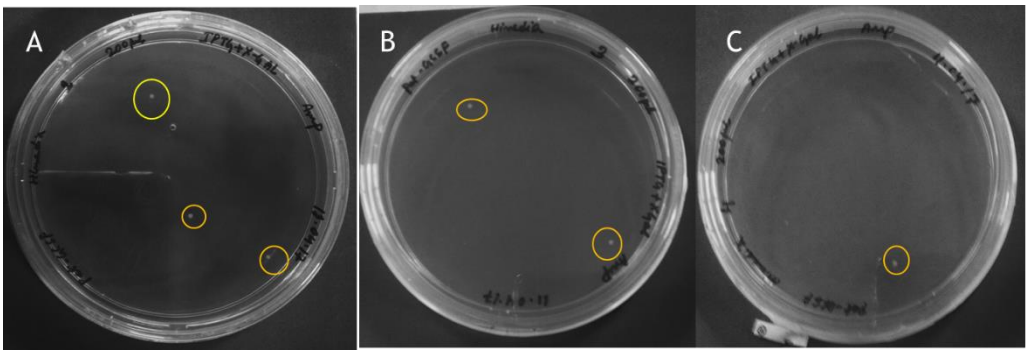


Figure 28: Transformation plates with ligated vector and insert A, B, C: Ligation Plates

Colonies were observed in the ligation plates along with no colony in the control plate. The colonies obtained in ligation plate were screened for presence of hG-CSF by PCR verification as shown in the figure 29.

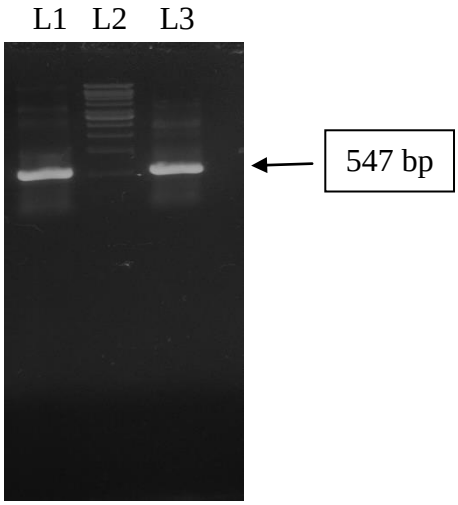


Figure 29: PCR verification of ligated colonies. Lane 1,3: PCR with hG-CSF specific primers.2:1kb Ladder

The two colonies screened were found to contain human G-CSF as a band of size 547bp was obtained after PCR amplification. Also some nonspecific bands were observed which

were in multiples of 500bp suggesting there were chances of tandem incorporation of the gene. Also plasmid was isolated and verified by running in agarose gel.

5.6 Cloning of pET14b-hGCSF construct in BL21DE3

After successful verification of ligation of hG-CSF to pET14b, the ligated plasmid was transformed to BL21DE3 expression host. The transformation plates are shown in figure 30.

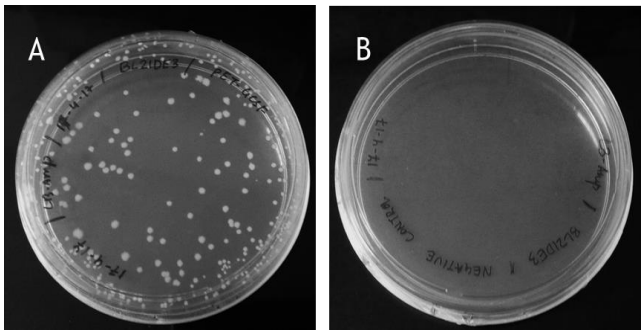


Figure 30: Transformation plates of pET14b-hGCSF construct in BL21DE3 A: Transformation plate B: Control Plate

Colonies obtained were inoculated and cultured overnight. Plasmid was isolated (using Charge Switch Pro Mini Prep Kit) and quantified to have A260/280 ratio of 1.9 and DNA conc of 35 ng/µl. The isolated plasmid DNA was verified for the presence of human G-CSF by PCR using human G-CSF specific primers. The result obtained is shown in the figure 31. All the plasmid gave PCR amplification for human GCSF.

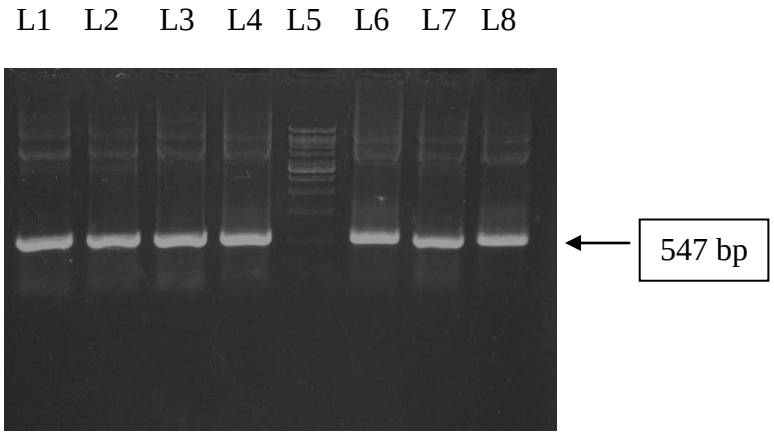


Figure 31: PCR verification of transformed colonies. Lane 1,2,3,4,6,7,8: PCR with GCSF specific primers.5:1kb Ladder

5.7 *In silico* analysis of hG-CSF with mutations and fusion partner for improved bioactivity

From the literature reviews it has been reported that N-terminal mutations have shown to increase the bioactivity of protein. With reference to this following five types of mutants were designed.

Sites	**2**	**4**	**5**	**6**	**18**
Wild	T	L	G	P	C
M1	A	T	Y	R	S
M2	A	T	F	R	S
M3	A	T	Y	K	S
M4	G	T	Y	R	S
M5	A	Y	Y	R	S

SWISS-MODEL was used for the homology modeling of the native and the mutant proteins. The protein sequences were given as input for the modeler and it generated models of the protein PyMOL was used to visualize the modeled proteins by giving pdb files of the mutant proteins.

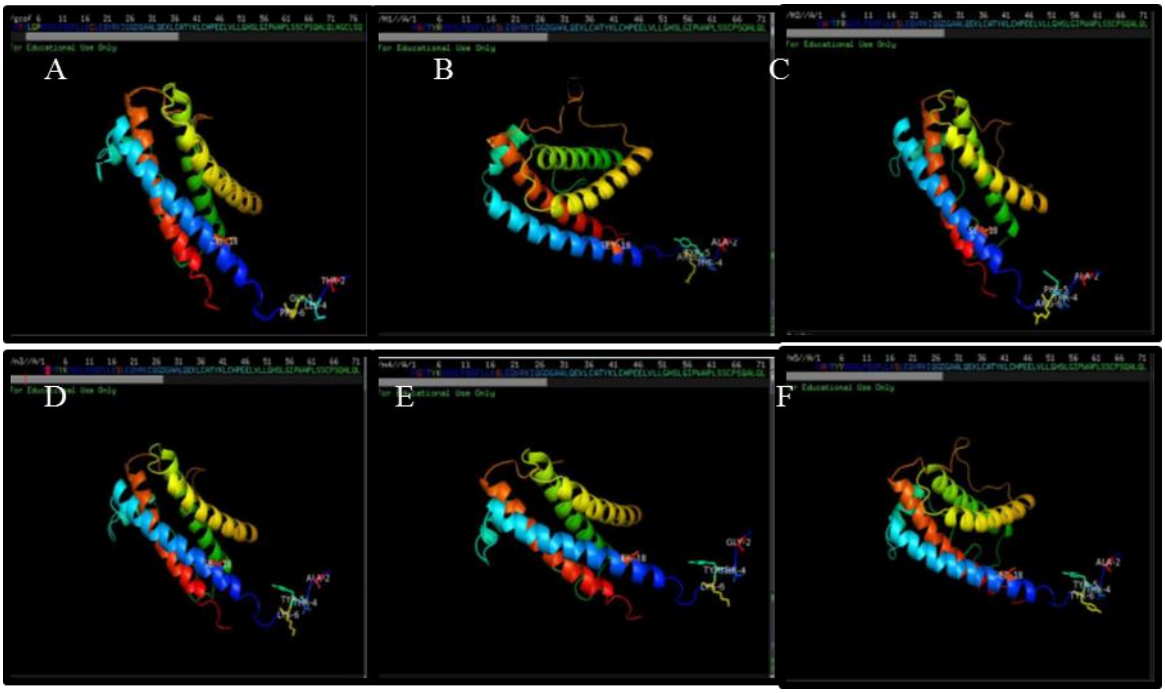


Figure 32: Proteins modeled using SWISS-MODEL and visualized using PyMOL. A: Native protein B:M1 mutant C:M2 mutant D:M3 mutant E:M4 mutant F:M5 mutant

The native GCSF showed 100% similarity with template 1gnc.1. A which had a description of granulocyte colony stimulating factor. Mutants M1, M2, M3, M4 and M5 showed 97.14,97.14,97.13,97.14 and 98.84% similarity respectively. Hex 8.0.0 was used to study the receptor-ligand interaction. GCSF receptor was docked with the native and all the five mutants.

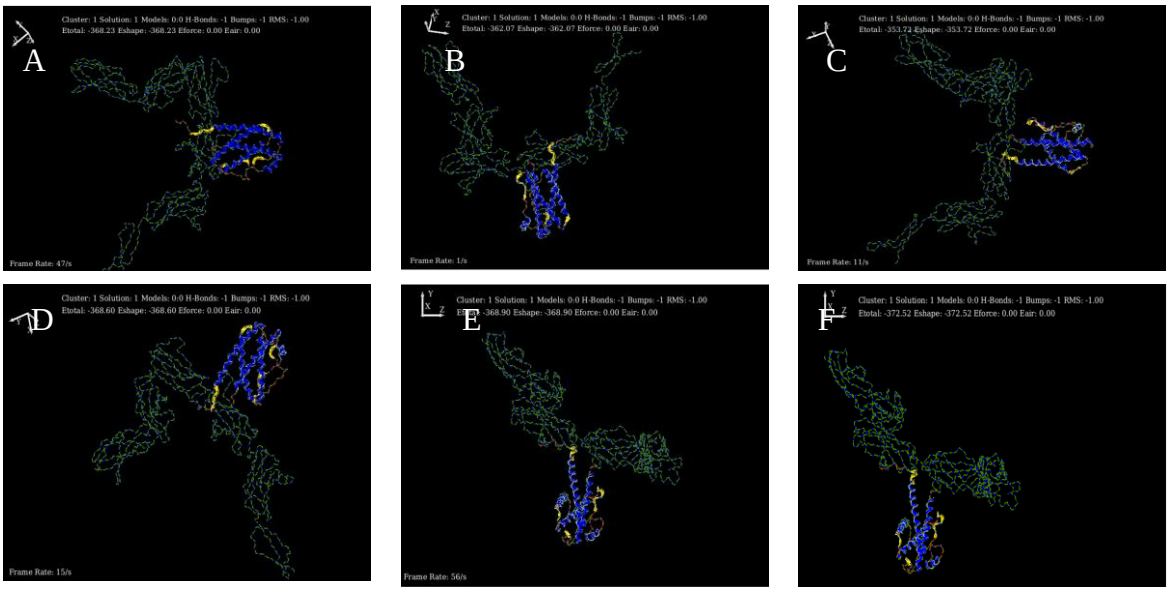


Figure 33: Receptor Protein interaction using Hex 8.0.0.A: Native protein B:M1 mutant C:M2mutant D:M3 mutant E:M4 mutantF:M5 mutant

Table 9: ETOTAL values for the various interactions

PROTEIN	ETOTAL(KJ/mol)
NATIVE	-368.23
M1	-362.07
M2	-353.72
M3	-368.6
M4	-368.9
M5	-372.52

ETotal values were used to compare the binding interactions between the proteins.Mutant 3,4 and 5 gave more ETOTAL value than the native protein inferring that all these mutants increased the binding affinity with M5 showing the highest affinity.M2 energy values were comparable to the native protein whereas M2 had a less binding affinity than the native one. Genome compiler and quick change site directed mutagenesis platform can be used to design the mutagenic primers.

In silico analysis of GCSF fusion protein was carried out to check the interaction of GCSF fusion protein with respective receptors and with themselves. Tertiary structure of the GCSF fusion proteins were analyzed using I-TASSER to find the interaction of fusion proteins with its respective receptor.

The three-dimensional structural of protein is important in providing insights to their molecular functions. The 3Dmodel of the GCSF/SCF fusion protein was generated using the I-TASSER online server along with their confidence score. Five models have generated by this server with C-Scores-3.41, 3.78, -4.05, -4.24, ¬-3.74 respectively. Among the 5models, model 1 has selected for further analysis as it contained the highest C-Score.

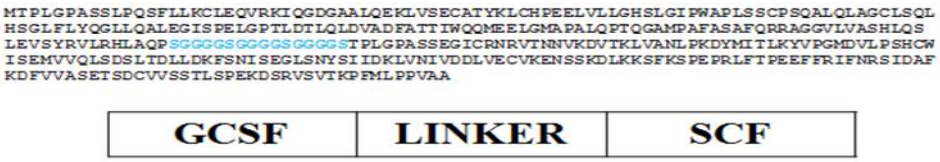


Figure 34: Sequence and schematic model which has shown the construct of GCSF and SCF bound together by the SGGGGSGGGGSGGGGS linker.

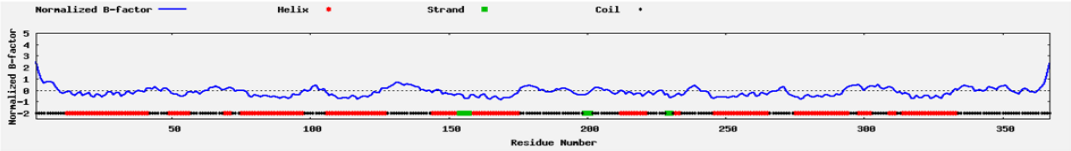


Figure 35: Graphical representation of secondary elements in chimeric GCSF-SCF protein

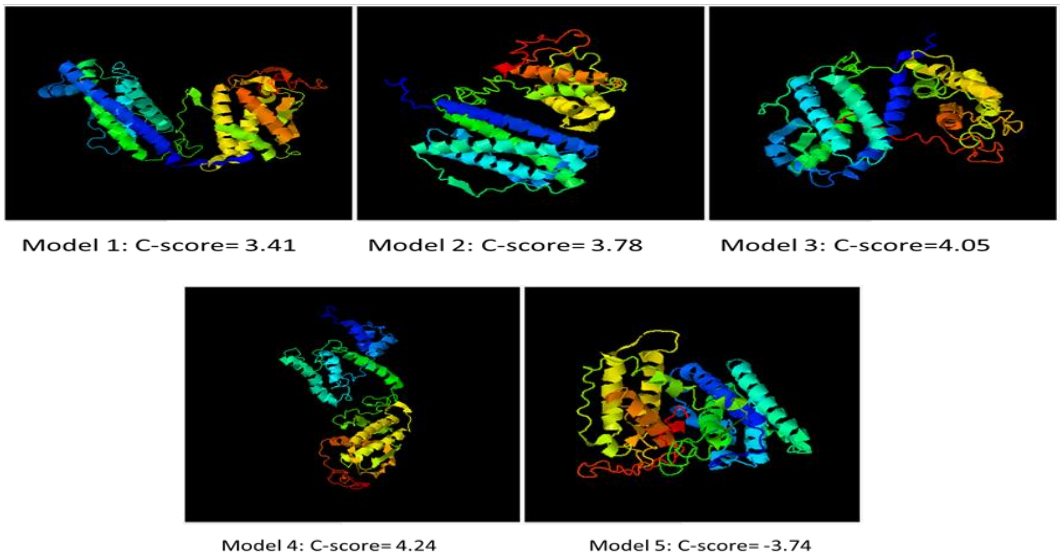
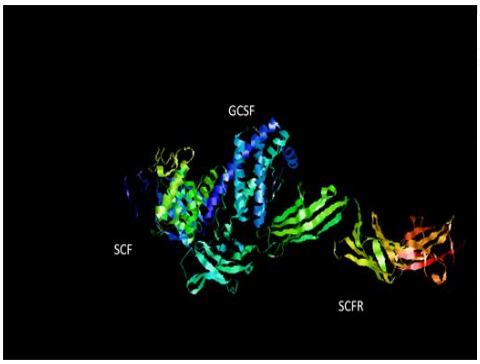


Figure 36: I-Tasser server was used to predict the 3D model of GCSF-SCF fusion protein

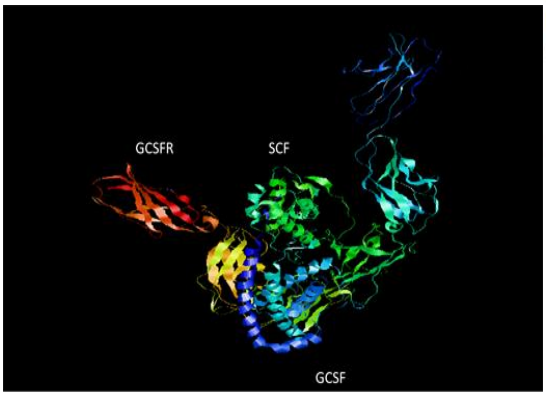
Ligand docking was performed using patch dock server. 3D Models of ligand (GCSF/SCF fusion protein) and receptors (GCSF receptor & SCF receptor) were uploaded in the software. Default parameters were used for the interaction studies. Global energy value was used to select the efficient docking process. More efficient the docking more negatives the global energy value.



Receptor: model.pdb | Ligand: model1.pdb | TransFil: fd_trans

Rank	Solution Number	Global Energy ↓	Attractive VdW	Repulsive VdW	ACE	HB	Structure show/hide
1	10	-9.94	-10.50	3.10	1.76	-2.83	☑
2	2	-6.98	-6.59	0.76	-2.83	-0.37	☑
3	7	-6.09	-13.28	12.83	3.22	-2.11	☐
4	5	-1.16	-3.84	0.32	1.28	-0.73	☐
5	8	1.62	-3.71	0.58	-1.17	0.00	☐
6	1	674.38	-30.22	886.10	7.27	-4.75	☐
7	3	850.18	-16.62	1082.89	7.21	-2.32	☐
8	6	978.06	-51.48	1332.29	1.13	-9.01	☐
9	4	1636.83	-59.78	2155.28	-2.90	-7.68	☐
10	9	6424.51	-99.85	8262.61	-23.21	-17.36	☐

Figure 37: Docking of GCSF-SCF with GCSFR using patchdock. To examine the protein ligand interactions, the models for ligand binding potency has predicted



Receptor: scfr.pdb | Ligand: model1.pdb | TransFile: fd_trans.txt

Rank	Solution Number	Global Energy ↓	Attractive VdW	Repulsive VdW	ACE	HB	Structure show/hide
1	10	-16.77	-24.25	18.81	4.74	-2.26	☑
2	9	9.63	-4.52	1.65	1.99	-0.44	☐
3	7	10.69	-2.46	1.50	1.79	-0.90	☐
4	3	19.84	-40.44	17.10	17.14	-3.45	☐
5	6	247.04	-17.03	294.47	12.25	-1.70	☐
6	2	317.30	-23.17	401.32	13.87	-5.18	☐
7	1	583.97	-43.46	746.71	12.58	-6.41	☐
8	5	983.03	-47.12	1289.04	11.27	-10.25	☐
9	4	1002.07	-21.24	1263.65	4.77	-2.74	☐
10	8	1636.92	-54.22	2204.31	-18.34	-5.45	☐

Figure 38: Docking of GCSF-SCF with GCSFR using patchdock. To examine the protein ligand interactions, the models for ligand binding potency has predicted.

Global energy value obtained from the docking of GCSF-SCF fusion protein with GCSFR and SCFR receptors were -9.94 and -16.77 respectively. Negative values of the docking process confirmed the higher affinity of the domains of fusion protein with its respective receptors. The result also confirmed that the fusion of SCF protein with GCSF protein do not affect the affinity of GCSF to its receptor. It was found that the SCF could be a good fusion partner of GCSF that enhance the bioactivity of GCSF.

5.8 Cloning of SCF in pUC57 cloning vector

The pUC57-SCF plasmid was transformed to HiPuRA DH5α Competent cells. Colonies were observed in the transformation plate as shown in the figure 39. No colonies were observed in the negative transformation plate.

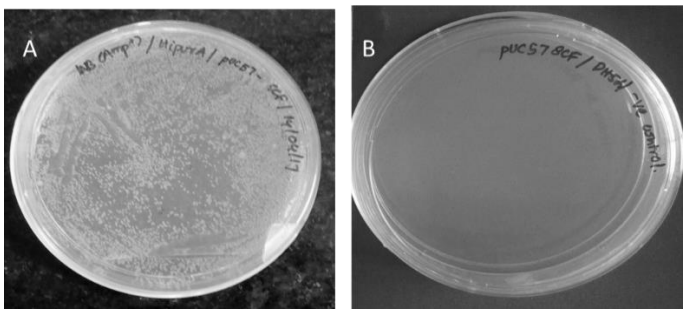


Figure 39: Transformation plate of pUC57-SCF to DH5α Competent Cells A: Transformation plate B: Negative control plate

After successful transformation, colonies were inoculated and grown overnight to proceed for isolation of plasmid. For verification of successful transformation, PCR Amplification of the isolated plasmid with SCF specific primers was carried out as shown in figure 40. Restriction Digestion of pET14b [Nde I & Xho I] was performed and the product was runned in agarose gel as shown in the figure 41.

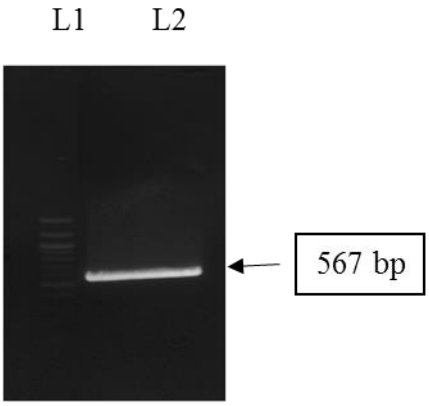


Figure 40: PCR amplification of human SCF with specific primers. Lane 1: 1kb Ladder 2: PCR with specific primer Band of size of around 567bp was observed in the lane loaded with PCR using SCF specific primers

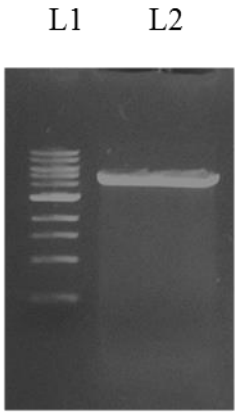


Figure 41: Restriction Digestion of pET14b with Nde I & Xho I.Lane 1:1kb Ladder 2:Restriction digested product.

Bands of appropriate size was obtained. Restriction Digestion of GCSF [Nde I and BamHI] and SCF [BamHI and XhoI] was further performed as shown in the figure 42.

L1 L2 L3

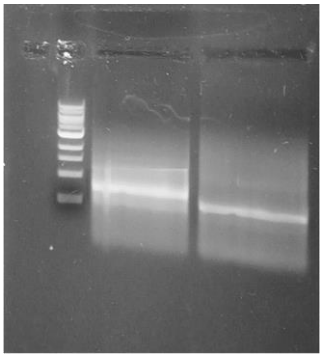


Figure 42: Restriction Digestion of pET14b with Nde I & Xho I. Lane 1:1kb Ladder 2: Restriction digested GCSF product 3: Restriction digested SCF product

5.9 Ligation of pET14b-hGCSF-SCF

Three-Way Ligation of the restriction digested pET14b-GCSF-SCF was performed by adding all the samples in 1:1:1 ratio and was transformed to to DH5α Competent Cells. The transformation plates are shown in the figure 43.

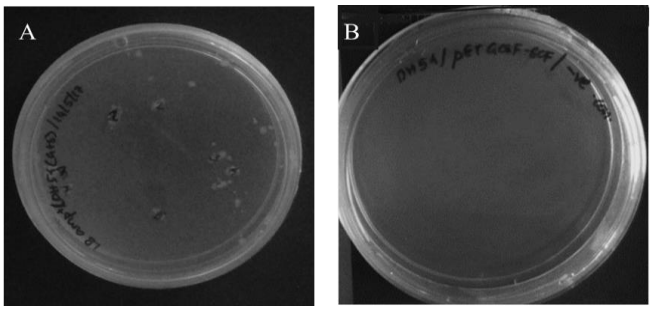


Figure 43: Transformation plates with ligated vector and insert A: Ligation Plate B: Control Plate

Colonies were obtained in the transformed plate whereas no colonies were obtained in control plate without plasmid. The colonies obtained were screened for the presence of GCSF and SCF by colony PCR. The colony PCR result is shown in figure 44.

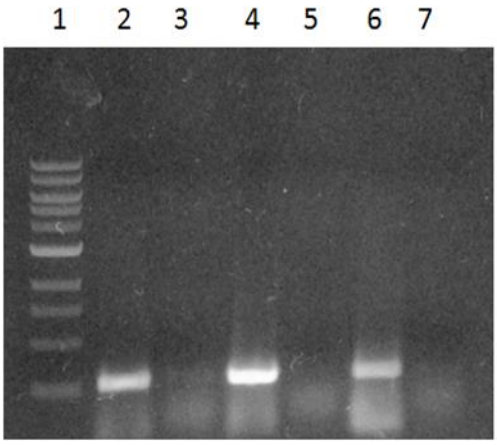


Figure 44: Colony PCR result. Lane 1:1 kb ladder 2: Colony 1with GCSF primer 3: Colony 1 with SCF primer 4: Colony 2 with SCF primer 5: Colony 3 with GCSF primer6: Colony 3 with SCF primer

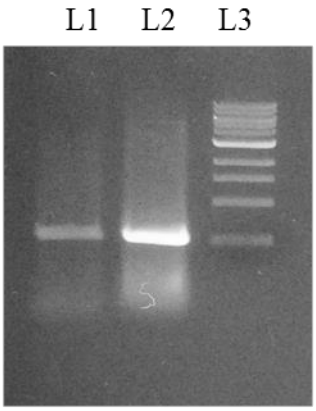


Figure 45: PCR verification of GCSF and SCF inserts in the isolated plasmid. Lane 1: 1- 1 kb ladder 2: PCR verification with both GCSF and SCF primers 3: PCR verification with SCF primer

The colonies which gave positive result for SCF and GCSF were chosen for further analysis. Positive colonies were inoculated and grown overnight for plasmid isolation. The isolated plasmid was PCR verified for the presence of GCSF and SCF by using specific primers as shown in figure 45. Both the primers gave band of specific size with no nonspecific bands confirming the presence of both SCF and GCSF in the plasmid.

5.10 Cloning of pET14b-hGCSF-SCF in BL21DE3

The plasmid was then transformed to BL21DE3 expression host as shown in figure 46.

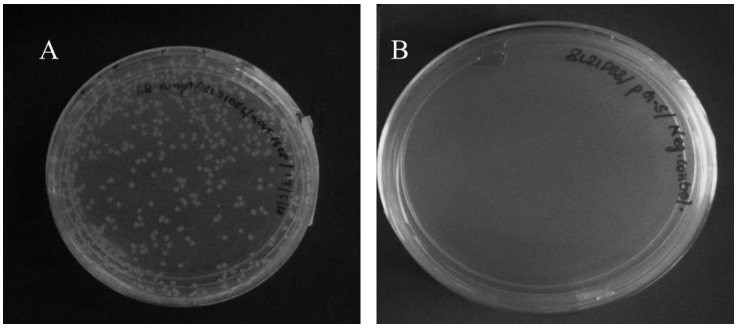


Figure 46: Transformation of plasmid to BL21DE3 competent cells A: Transformation plate B: Negative control plate

5.11 Expression of human G-CSF

1mM of IPTG was used to induce the overnight grown culture when the O.D. reached 0.5. O.D.600nm spectrophotometric readings of the induced and uninduced culture was recorded hourly. 0,4,6,8 and overnight grown induced and uninduced culture was harvested and proceeded for sonication and protein expression.

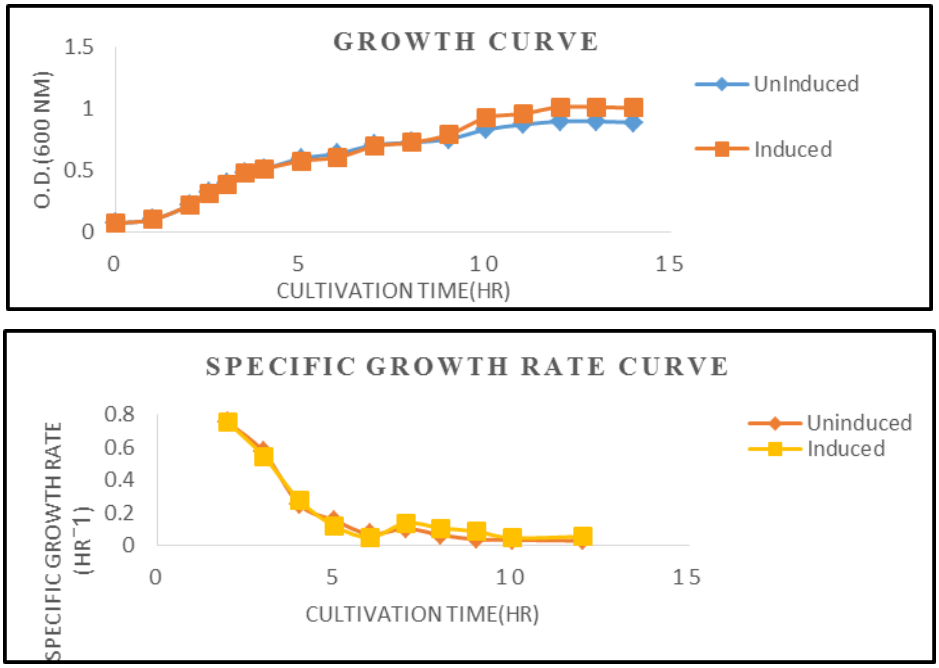


Figure 47: Growth curve and specific growth rate curve of induced and uninduced culture.

Further ethanol (%) was added to the culture to increase the biomass of the culture.The growth curve was anaysed for 0%,1%,2% and 3% ethanol addition and the graphs were plotted for the same as shown in figure 49.

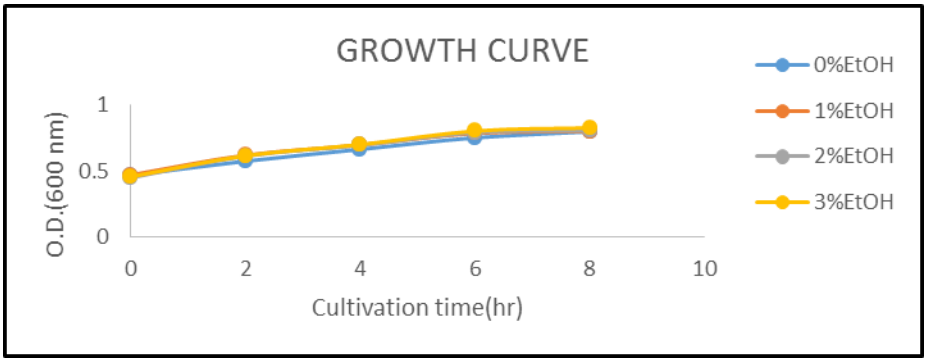


Figure 48: Growth curve at different ethanol concentration.

Both the growth curve and specific growth rate curve were higher for induced culture than uninduced culture.3% ethanol gave the highest growth upon analysis of effect of ethanol
on culture conditions. The SDS-PAGE results for the same is shown in the figure 49.

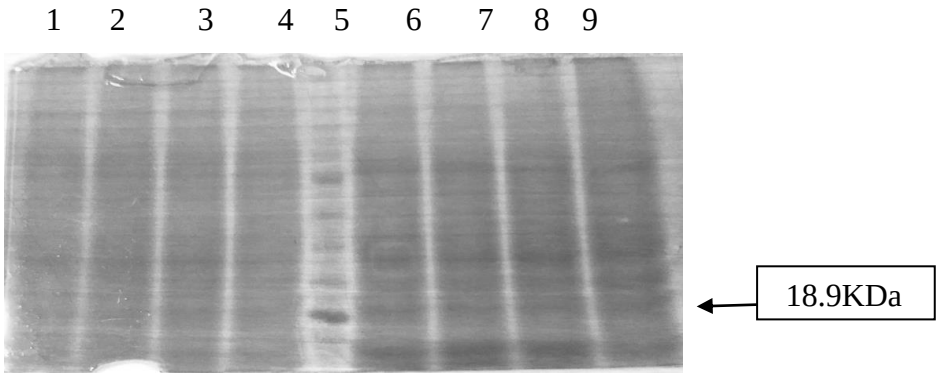


Figure 49: SDS-PAGE result of induced and uninduced culture of GCSF. Lane 1-4: Uninduced samples at 4,6,8 and overnight 5: Protein Ladder.6-9: Induced samples at 4,6,8 and overnight cultivation time.

5.12 Expression of human G-CSF-SCF

1mM of IPTG was used to induce the overnight grown culture containing fusion protein when the O.D. reached 0.5. O.D.600 nm spectrophotometric readings of the induced and uninduced culture was recorded hourly. 0,4,6,8 and overnight grown induced and uninduced culture was harvested and proceeded for sonication and protein expression. The SDS-PAGE result for the same is shown in figure 50.

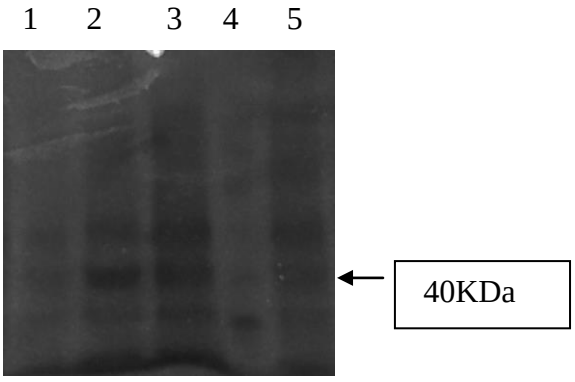


Figure 50: SDS-PAGE result of induced and uninduced culture of GCSF-SCF.Lane1: Uninduced,2: Induced (6 hr),3: Induced (8hr) 4: Ladder 5: Induced (12 hr)

Protein band of 40 KDa was obtained as analyzed from the SDS-PAGE. No band was observed in the uninduced sample. However, for further confirmation western blotting with specific antibody is required.

Chapter 6

Conclusion

Total RNA was isolated from U-87 cell lines and umbilical cord blood, converted to cDNA and amplified with hG-CSF specific primers and hG-CSF specific EEB primers. The DNA obtained was then cloned into a TOPO vector and pUC57. Transformation of pUC57-hG-CSF was performed and the colonies obtained were verified for the presence of hG-CSF using PCR. The expression vector pET14b was chosen for the expression of the protein. After transforming and culturing the vector in E. coli cells it was restriction digested with Nde I and Bam HI for ligation with the digested hG-CSF. Ligated colonies were screened for the presence of human GCSF and then transformed to BL21DE3 expression host and PCR verified for presence of hG-CSF.

Mutations in N-terminus of hG-CSF have shown to increase the bioactivity. Keeping this in mind five mutants were designed aiming to increase the bioactivity of the wild-type protein. All the mutants and the wild type were modelled and then their docking with the receptor was studied. Energy values were found to be more negative than the wild type in three of the five mutants. Genome compiler and Quick change primer design platform can be used to design the mutagenic primers to carry out the modifications. In order to check the interaction of fusion partner with GCSF protein in terms of receptor binding and folding in the fusion protein conformation, *in silico* analysis was carried out using I-TASSER and patch dock sever respectively. It was found that the global energy values of docking of hG-CSF-SCF fusion protein with GCSFR and SCFR were -9.94 and -16.77 respectively. This indicates the high affinity binding of the domains of hG-CSF-SCF fusion protein with its corresponding receptor. The visualization of docking result also suggested that the binding of one domain of the fusion protein in the binding site of receptor was not hindered by another domain. These findings suggested that the SCF could be a correct fusion partner of the GCSF protein.

SCF was successfully cloned in pUC57 vector and a three-way ligation was carried out to fuse SCF with hG-CSF and pET14b. The colonies obtained was PCR verified to contain both hG-CSF and SCF and then transformed to DE3 competent cells. Growth profiles of

induced and uninduced culture were determined and protein expression was analysed. Induction for 8h gave the maximum protein expression as analysed by Image J software. IPTG concentration and ethanol (%) was also optimized for analysing increased protein expression.

The IPTG concentration of 1mM along with 8hr of induction in the presence of 3% of ethanol was optimized for higher protein expression. Both hG-CSF and hG-CSF-SCF fusion partner was successfully expressed as analysed from SDS-PAGE. However, for further confirmation western blotting with hG-CSF specific antibody is required. Furthermore, other bioprocess parameters need to be optimized for increasing the protein expression along with purification and bioactivity analysis.

References

1. Anderson, Dirk M., Douglas E. Williams, R. Tushinski, S. Gimpel, J. Eisenman, L. A. Cannizzaro, M. Aronson et al. "Alternate splicing of mRNAs encoding human mast cell growth factor and localization of the gene to chromosome 12q22-q24." *Cell Growth Differ* 2, no. 8 (1991): 373-378.
2. Anderson, Dirk M., Stewart D. Lyman, Allison Baird, Janis M. Wignall, June Eisenman, Charles Rauch, Carl J. March et al. "Molecular cloning of mast cell growth factor, a hematopoietin that is active in both membranes bound and soluble forms." *Cell* 63, no. 1 (1990): 235-243.
3. Andrews, Robert G., Robert A. Briddell, Glenn H. Knitter, Scott D. Rowley, Frederick R. Appelbaum, and Ian K. McNiece. "Rapid engraftment by peripheral blood progenitor cells mobilized by recombinant human stem cell factor and recombinant human granulocyte colony-stimulating factor in nonhuman primates." *Blood* 85, no. 1 (1995): 15-20.
4. Arakawa, T., D. A. Yphantis, J. W. Lary, L. O. Narhi, H. S. Lu, S. J. Prestrelski, C. L. Clogston, K. M. Zsebo, E. A. Mendiaz, and J. Wypych. "Glycosylated and unglycosylated recombinant-derived human stem cell factors are dimeric and have extensive regular secondary structure." *Journal of Biological Chemistry* 266, no. 28 (1991): 18942-18948.
5. Aritomi, Masaharu, Naoki Kunishima, Tomoyuki Okamoto, Ryota Kuroki, Yoshimi Ota, and Kosuke Morikawa. "Atomic structure of the GCSF–receptor complex showing a new cytokine–receptor recognition scheme." *Nature* 401, no. 6754 (1999): 713-717.
6. Babaeipour, Valiollah, Mahdi Pesaran Haji Abbas, Zahra Sahebnazar, and Reza Alizadeh. "Enhancement of human granulocyte-colony stimulating factor production in recombinant E. coli using batch cultivation." *Bioprocess and biosystems engineering* 33, no. 5 (2010): 591-598.
7. Bai, Yun, David K. Ann, and Wei-Chiang Shen. "Recombinant granulocyte colony-stimulating factor-transferrin fusion protein as an oral myelopoietic agent." *Proceedings of the National Academy of Sciences of the United States of America* 102, no. 20 (2005): 7292-7296.
8. Basu, Sunanda, Ashley Dunn, and Alister Ward. "G-CSF: function and modes of action (Review)." *International journal of molecular medicine* 10, no. 1 (2002): 3-10.
9. Bishop, Barney, Debbie C. Koay, Alan C. Sartorelli, and Lynne Regan. "Reengineering granulocyte colony-stimulating factor for enhanced stability." *Journal of Biological Chemistry* 276, no. 36 (2001): 33465-33470.
10. Briddell, R. A., C. A. Hartley, K. A. Smith, and I. K. McNiece. "Recombinant rat stem cell factor synergizes with recombinant human granulocyte colony-stimulating factor in vivo in mice to mobilize peripheral blood progenitor cells that have enhanced repopulating potential." *Blood* 82, no. 6 (1993): 1720-1723.
11. Broudy, Virginia C. "Stem cell factor and hematopoiesis." *Blood* 90, no. 4 (1997): 1345-1364.

12. Chen, Rachel. "Bacterial expression systems for recombinant protein production: E. coli and beyond." *Biotechnology advances* 30, no. 5 (2012): 1102-1107.
13. Chen, Tao, Yuhui Zang, Jie Zhu, Haiqin Lu, Junhai Han, and Junchuan Qin. "Expression of a novel recombinant stem cell factor/macrophage colony-stimulating factor fusion protein in baculovirus-infected insect cells." *Protein expression and purification* 41, no. 2 (2005): 402-408.
14. Chung, Hee-Kyoung, Sung-Woo Kim, Sung-June Byun, Eun-Mi Ko, Hak-Jae Chung, Jae-Seok Woo, Jae-Gyu Yoo et al. "Enhanced biological effects of Phe140Asn, a novel human granulocyte colony-stimulating factor mutant, on HL60 cells." *BMB reports* 44, no. 10 (2011): 686-691.
15. Cooper, Katy L., Jason Madan, Sophie Whyte, Matt D. Stevenson, and Ron L. Akehurst. "Granulocyte colony-stimulating factors for febrile neutropenia prophylaxis following chemotherapy: systematic review and meta-analysis." *BMC cancer* 11, no. 1 (2011): 404.
16. Copeland, Neal G., Debra J. Gilbert, Brian C. Cho, Peter J. Donovan, Nancy A. Jenkins, David Cosman, Dirk Anderson, Stewart D. Lyman, and Douglas E. Williams. "Mast cell growth factor maps near the steel locus on mouse chromosome 10 and is deleted in a number of steel alleles." *Cell* 63, no. 1 (1990): 175-183.
17. Crawford, Jeffrey, David C. Dale, and Gary H. Lyman. "Chemotherapy-induced neutropenia." *Cancer* 100, no. 2 (2004): 228-237.
18. De Revel, Thierry, Frederick R. Appelbaum, Rainer Storb, Friedrich Schuening, Richard Nash, Joachim Deeg, Ian McNiece, Robert Andrews, and Theodore Graham. "Effects of granulocyte colony-stimulating factor and stem cell factor, alone and in combination, on the mobilization of peripheral blood cells that engraft lethally irradiated dogs." *Blood* 83, no. 12 (1994): 3795-3799.
19. Dingermann, Theo. "Recombinant therapeutic proteins: production platforms and challenges." *Biotechnology journal* 3, no. 1 (2008): 90-97.
20. Do, Bich Hang, Han-Bong Ryu, Phuong Hoang, Bon-Kyung Koo, and Han Choe. "Soluble prokaryotic overexpression and purification of bioactive human granulocyte colony-stimulating factor by maltose binding protein and protein disulfide isomerase." *PloS one* 9, no. 3 (2014): e89906.
21. Duarte, Rafael F., and David A. Frank. "The synergy between stem cell factor (SCF) and granulocyte colony-stimulating factor (G-CSF): molecular basis and clinical relevance." *Leukemia & lymphoma* 43, no. 6 (2002): 1179-1187.
22. Facon, Thierry, Jean-Luc Harousseau, Frédéric Maloisel, Michel Attal, Jesus Odriozola, Adrian Alegre, Wilfried Schroyens et al. "Stem cell factor in combination with filgrastim after chemotherapy improves peripheral blood progenitor cell yield and reduces apheresis requirements in multiple myeloma patients: a randomized, controlled trial." *Blood* 94, no. 4 (1999): 1218-1225.
23. Fallah, M. J., B. Akbari, A. R. Saeedinia, M. O. H. S. E. N. Karimi, M. Vaez, M. Zeinoddini, M. Soleimani, and N. Maghsoudi. "Overexpression of recombinant human granulocyte colony-stimulating factor in E. coli." *Iranian Journal of Medical Sciences* 28, no. 3 (2015): 131-134.

24. Fujii, Isao, Yoshitomo Nagahara, Motoo Yamasaki, Yoshiharu Yokoo, Seiga Itoh, and Noriaki Hirayama. "Structure of KW-2228, a tailored human granulocyte colony-stimulating factor with enhanced biological activity and stability." *FEBS letters* 410, no. 2-3 (1997): 131-135.
25. Gomes, F. R., A. C. Maluenda, J. O. Tápias, F. L. S. Oliveira, L. C. Sá-Rocha, E. Carvalho, and P. L. Ho. "Expression of recombinant human mutant granulocyte colony stimulating factor (Nartograstim) in Escherichia coli." *World Journal of Microbiology and Biotechnology* 28, no. 7 (2012): 2593-2600.
26. Huang, Eric, Karl Nocka, David R. Beier, Tang-Yan Chu, Jochen Buck, Hans-Werner Lahm, Daniel Wellner, Philip Leder, and Peter Besmer. "The hematopoietic growth factor KL is encoded by the SI locus and is the ligand of the c-kit receptor, the gene product of the W locus." *Cell* 63, no. 1 (1990): 225-233.
27. Hill, Christopher P., Timothy D. Osslund, and David Eisenberg. "The structure of granulocyte-colony-stimulating factor and its relationship to other growth factors." *Proceedings of the National Academy of Sciences* 90, no. 11 (1993): 5167-5171.
28. Ishikawa, Masaharu, Hiroshi Iijima, Rika Satake-Ishikawa, Haruhiko Tsumura, Akihiro Iwamatsu, Toshihiko Kadoya, Yoshihiro Shimada et al. "The Substitution of Cysteine 17 of Recombinant HumanG-CSF with Alanine Greatly Enhanced its Stability." *Cell structure and function* 17, no. 1 (1992): 61-65.
29. Keefer, C. L. "Production of bioproducts through the use of transgenic animal models." *Animal Reproduction Science* 82 (2004): 5-12
30. Kubota, Naoki, Tetsuro Orita, Kunihiro Hattori, Masayoshi Oh-eda, Norimichi Ochi, and Tatsumi Yamazaki. "Structural characterization of natural and recombinant human granulocyte colony-stimulating factors." *The Journal of Biochemistry* 107, no. 3 (1990): 486-492.
31. Lu, Hsieng S., Patricia R. Fausset, Linda O. Narhi, Thomas Horan, Kyoko Shinagawa, Grant Shimamoto, and Thomas C. Boone. "Chemical modification and site-directed mutagenesis of methionine residues in recombinant human granulocyte colony-stimulating factor: effect on stability and biological activity." *Archives of biochemistry and biophysics* 362, no. 1 (1999): 1-11.
32. Maniatis, Tom, Edward F. Fritsch, and Joseph Sambrook. *Molecular cloning: a laboratory manual*. Vol. 545. Cold Spring Harbor, NY: Cold Spring harbor laboratory, 1982.
33. Martin, F. H. "Suggs SV, Langley KE, Lu HS, Ting J, Okino KH, Morris CF, McNiece IK, Jacobsen." *FW, Mendiaz EA, Birkett NC, Smith KA, Johnson MJ, Parker VP, Flores JC, Pate1 AC, Fisher EF, Erjavek HO, Herrera CJ, Wypych J, Sachdev RK, Pope JA, Leslie I, Wen D, Lin CH, Cupples RL, Zsebo KM: Primary structure and functional expression of rat and human stem cell factor DNAs. Cell* 63, no. 203 (1990): 1991.
34. Maullu, Carlo, Domenico Raimondo, Francesca Caboi, Alejandro Giorgetti, Mauro Sergi, Maria Valentini, Giancarlo Tonon, and Anna Tramontano. "Site-directed enzymatic PEGylation of the human granulocyte colony-stimulating factor." *FEBS journal* 276, no. 22 (2009): 6741-6750.

35. Metcalf, Donald. "Hematopoietic regulators: redundancy or subtlety?." *BLOOD-NEW YORK*- 82 (1993): 3515-3515.
36. Meuer, Katrin, Claudia Pitzer, Peter Teismann, Carola Krüger, Bettina Göricke, Rico Laage, Paul Lingor et al. "Granulocyte-colony stimulating factor is neuroprotective in a model of Parkinson's disease." *Journal of neurochemistry* 97, no. 3 (2006): 675-686.
37. Moskowitz, Craig H., Patrick Stiff, Michael S. Gordon, Ian McNiece, Anthony D. Ho, John J. Costa, E. Randolf Broun et al. "Recombinant methionyl human stem cell factor and filgrastim for peripheral blood progenitor cell mobilization and transplantation in non-Hodgkin's lymphoma patients—results of a phase I/II trial." *Blood* 89, no. 9 (1997): 3136-3147.
38. Nagata, Shigekazu, Masayuki Tsuchiya, Shigetaka Asano, O. Yamamoto, Y. Hirata, N. Kubota, M. Oheda, H. Nomura, and T. Yamazaki. "The chromosomal gene structure and two mRNAs for human granulocyte colony-stimulating factor." *The EMBO Journal* 5, no. 3 (1986): 575.
39. Nagata, Shigekazu, Masayuki Tsuchiya, Shigetaka Asano, Yoshito Kaziro, Tatsumi Yamazaki, Osami Yamamoto, Yuichi Hirata et al. "Molecular cloning and expression of cDNA for human granulocyte colony-stimulating factor." *Nature* 319, no. 6052 (1986): 415-418.
40. Nagata, Shigekazu. "Gene structure and function of granulocyte colony-stimulating factor." *Bioessays* 10, no. 4 (1989): 113-117
41. Petit, I., Szyper-Kravitz, M., Nagler, A., Lahav, M., Peled, A., Habler, L., Ponomaryov, T., Taichman, R.S., Arenzana-Seisdedos, F., Fujii, N. and Sandbank, J., 2002. G-CSF induces stem cell mobilization by decreasing bone marrow SDF-1 and up-regulating CXCR4. *Nature immunology*, *3*(7), pp.687-694.
42. Rao, Dasari V. Krishna, Joginapally V. Rao, Mangamoori L. Narasu, and Adibhatla Kali S. Bhujanga Rao. "Optimization of the AT-content of codons immediately downstream of the initiation codon and evaluation of culture conditions for high-level expression of recombinant human G-CSF in Escherichia coli." *Molecular biotechnology* 38, no. 3 (2008): 221-232.
43. Ria, Roberto, Antonia Reale, Assunta Melaccio, Vito Racanelli, Franco Dammacco, and Angelo Vacca. "Filgrastim, lenograstim and pegfilgrastim in the mobilization of peripheral blood progenitor cells in patients with lymphoproliferative malignancies." *Clinical and experimental medicine* 15, no. 2 (2015): 145-150.
44. Shpall, Elizabeth J., Catherine A. Wheeler, Stewart A. Turner, Saul Yanovich, Randy A. Brown, Andrew L. Pecora, Thomas C. Shea et al. "A randomized phase 3 study of peripheral blood progenitor cell mobilization with stem cell factor and filgrastim in high-risk breast cancer patients." *Blood* 93, no. 8 (1999): 2491-2501.
45. Shuman, Stewart. "Recombination mediated by vaccinia virus DNA topoisomerase I in Escherichia coli is sequence specific." *Proceedings of the National Academy of Sciences* 88, no. 22 (1991): 10104-10108.
46. Shuman, Stewart. "Recombination mediated by vaccinia virus DNA topoisomerase I in Escherichia coli is sequence specific." *Proceedings of the National Academy of Sciences* 88, no. 22 (1991): 10104-10108.

47. Stanley, Edouard, Graham J. Lieschke, Dianne Grail, Donald Metcalf, George Hodgson, J. A. Gall, Darryl W. Maher, Jonathan Cebon, Vincent Sinickas, and Ashley R. Dunn. "Granulocyte/macrophage colony-stimulating factor-deficient mice show no major perturbation of hematopoiesis but develop a characteristic pulmonary pathology." *Proceedings of the National Academy of Sciences* 91, no. 12 (1994): 5592-5596.

48. Tamada, Taro, Eijiro Honjo, Yoshitake Maeda, Tomoyuki Okamoto, Matsujiro Ishibashi, Masao Tokunaga, and Ryota Kuroki. "Homodimeric cross-over structure of the human granulocyte colony-stimulating factor (GCSF) receptor signaling complex." *Proceedings of the National Academy of Sciences of the United States of America* 103, no. 9 (2006): 3135-3140.
49. Vanz, Ana LS, Gaby Renard, Mario S. Palma, Jocelei M. Chies, Sérgio L. Dalmora, Luiz A. Basso, and Diógenes S. Santos. "Human granulocyte colony stimulating factor (hG-CSF): cloning, overexpression, purification and characterization." *Microbial Cell Factories* 7, no. 1 (2008): 13.
50. Volberding, Paul A., Alexandra M. Levine, Douglas Dieterich, Donna Mildvan, Ronald Mitsuyasu, Michael Saag, and Anemia in HIV Working Group. "Anemia in HIV infection: clinical impact and evidence-based management strategies." *Clinical infectious diseases* 38, no. 10 (2004): 1454-1463.
51. Wen, Xiaofang, Yiliang Wu, Yefei Wang, Zhiyu Yang, Min Fan, Yujiao Wang, Xiaochun Fang, and You Lu. "Mutant G-CSF fusion protein, and preparation and use thereof." U.S. Patent 8,785,597, issued July 22, 2014.
52. Wittman, Brenda, John Horan, and Gary H. Lyman. "Prophylactic colony-stimulating factors in children receiving myelosuppressive chemotherapy: a meta-analysis of randomized controlled trials." *Cancer treatment reviews* 32, no. 4 (2006): 289-303.
53. Zhao, Shuqiang, Yu Zhang, Hong Tian, Xiaofei Chen, Di Cai, Wenbing Yao, and Xiangdong Gao. "Extending the serum half-life of G-CSF via fusion with the domain III of human serum albumin." *BioMed research international* 2013 (2013).
54. Zsebo, Krisztina M., David A. Williams, Edwin N. Geissler, Virginia C. Broudy, Francis H. Martin, Harry L. Atkins, Rou-Yin Hsu et al. "Stem cell factor is encoded at the SI locus of the mouse and is the ligand for the c-kit tyrosine kinase receptor." *Cell* 63, no. 1 (1990): 213-224.
55. Zsebo, Krisztina M., Jette Wypych, Ian K. McNiece, Hsieng S. Lu, Kent A. Smith, Subash B. Karkare, Raj K. Sachdev et al. "Identification, purification, and biological characterization of hematopoietic stem cell factor from buffalo rat liver-conditioned medium." *Cell* 63, no. 1 (1990): 195-201.